

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 6, 2004, 05:04:20 ; Search time 5952 Seconds

(without alignments)
11649.863 Million cell updates/sec

Title: US-10-074-744-4

Perfect score: 2322
Sequence: 1 aagctttctaatltaacct.....gtcgtatcgtggtgcacaca 2322

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estinu:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	635.2	27.4	722	10 AW983873	AW983873 RCI-HN000
2	628.6	27.1	649	10 AW983877	AW983877 RCI-HN000
3	602.8	26.0	663	10 AW983875	AW983875 RCI-HN000
4	497.6	21.4	547	10 AW984211	AW984211 PM2-HN000

5	400.8	17.3	446	10 BG010614	BG010614 CM4-GN029
6	314	13.5	336	14 R04703	R04703 pk27g02.r1
7	309.6	13.3	365	10 BF737769	BF737769 CM4-KT003
8	238.6	10.3	262	10 AW984213	AW984213 PM2-HN000
9	205.8	8.9	254	10 BF773183	BF773183 CM4-IT004
10	202.8	8.7	206	10 BF348952	BF348952 MRI-DT005
11	66.2	2.9	1101	29 CNS0039G	AL063921 Drosophila
12	65.8	2.8	1101	29 CNS0039G	AL063921 Drosophila
13	65.4	2.8	556	10 BF755979	BF755979 BB755979
14	61	2.6	994	13 BX414650	BX414650 BX414650
15	59.8	2.6	961	29 CNS008HT	AL051882 Drosophila
16	56.6	2.4	1201	13 BX355654	BX355654 BX355654
17	55.8	2.4	1146	29 CNS021G2	AL176843 Tetraodon
18	54	2.3	1201	13 BX461310	BX461310 BX461310
19	53.8	2.3	1091	13 BX424950	BX424950 BX424950
20	53.6	2.3	1124	13 BX436282	BX436282 BX436282
21	52.8	2.3	1101	29 CNS0106X	AL098595 Drosophila
22	52.4	2.3	1204	29 CNS016E2	AL106628 Drosophila
23	51.8	2.2	637	29 CNS036CC	AL229845 Tetraodon
24	51.2	2.2	897	29 CNS015FK	AL105386 Drosophila
25	51	2.2	1201	9 AL565455	AL565455 AL565455
26	50.8	2.2	598	12 BJ435904	BJ435904 BJ435904
27	50.8	2.2	1044	13 BX415231	BX415231 BX415231
28	50.6	2.2	1101	29 CNS0176D	AL107647 Drosophila
29	50.6	2.2	1101	29 CNS0182P	AL108811 Drosophila
30	50.6	2.2	1173	13 BX443524	BX443524 BX443524
31	50.4	2.2	928	29 CNS00DKY	AL071865 Drosophila
32	50.2	2.2	1199	13 BX375702	BX375702 BX375702
33	50	2.2	1056	13 BX415058	BX415058 BX415058
34	49.8	2.1	1201	13 BX361152	BX361152 BX361152
35	49.4	2.1	697	28 BZ012209	BZ012209 ce160g03
36	49.4	2.1	1201	13 BX396546	BX396546 BX396546
37	49.2	2.1	987	29 CNS014PQ	AL104456 Drosophila
38	49.2	2.1	1099	13 BX456575	BX456575 BX456575
39	49	2.1	1200	13 BX415878	BX415878 BX415878
40	48.8	2.1	1101	29 CNS00B3U	AL056389 Drosophila
41	48.6	2.1	1092	29 CNS020K7	AL175696 Tetraodon
42	48.6	2.1	1201	9 AL559324	AL559324 AL559324
43	48.4	2.1	1101	29 CNS0006J	AL062049 Drosophila
44	48.2	2.1	1201	13 BX363491	BX363491 BX363491
45	48.2	2.1	1201	13 BX386038	BX386038 BX386038

ALIGNMENTS

RESULT 1
LOCUS AW983873 722 bp mRNA linear EST 02-JUN-2000
DEFINITION RCI-HN0003-220300-021-c03 HN0003 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW983873
VERSION AW983873.1 GI:8175467
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 722)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjowski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=at2=RC1-HN0003-220300-021-c03&t3=2000-03-22&t4=1)

Seq primer: puc 18 forward
High quality sequence start: 40
High quality sequence stop: 685.

FEATURES

source

location/Qualifiers

1..722

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="HN0003"

/note="Organ: head_normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 27.4%; Score 635.2; DB 10; Length 722;

Best Local Similarity 98.6%; Pred. No. 2.3e-137;

Matches 651; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 6 TTCTTAATTTAACCTTTGTGAGGTACCAACTACTAAGGTTGTAGGCTCAAGAGGGTGTG 65

Db 28 TTTCTAATTTAAACC-TTGTGAGGTACCAACTACTAAGGTTGTAGGCTCAAGAGGGTGTG 86

QY 66 TCCTGTGCTAGGTAATAAAGTGAAGTGGAGTCTTAAATTTCTATATTGTTGTTCTTCT 125

Db 87 TCCTGTGCTAGGTAATAAAGTGAAGTGGAGTCTTAAATTTCTATATTGTTGTTCTTCT 146

QY 126 GCAAAAAAGTGGGAAGTGAATGAATTTATTTCTAATTTATCTGCATCATACCTT 185

Db 147 GCAAAAAAGTGGGAAGTGAATGAATTTATTTCTAATTTATCTGCATCATACCTT 206

QY 186 CCGAGCATTTATTAAGCATTTCCGCTAATAAGTTCTCGCTGGAAGAGTAGTTTTCATTG 245

Db 207 CCGAGCATTTATTAAGCATTTCCGCTAATAAGTTCTCGCTGGAAGAGTAGTTTTCATTG 266

QY 246 TACTTTACCTTCATCTCTGTTCAATATCATCGCTTTTAAACGGTTCGACCTTCTAATCC 305

Db 267 TACTTTACCTTCATCTCTGTTCAATATCATCGCTTTTAAACGGTTCGACCTTCTAATCC 326

QY 306 TATCTGACCATTAATATTTTGAATGGTTTCATAGAAGAGCTCTGAATCAACGGACTTG 365

Db 327 TATCTGACCATTAATATTTTGAATGGTTTCATAGAAGAGCTCTGAATCAACGGACTTG 386

QY 366 CGATAATAAGTGGTGTATCCAGAAATTTGTCACTTCAAGTAAACACCTCAGAGTTAA 425

Db 387 CGATAATAAGTGGTGTATCCAGAAATTTGTCACTTCAAGTAAACACCTCAGAGTTAA 446

QY 426 AACACCTAAGTCTTCACCGAATGTCTCAATATCCGACGGAATAATTTATTGCTTCTCT 485

Db 447 AACACCTAAGTCTTCACCGAATGTCTCAATATCCGACGGAATAATTTATTGCTTCTCT 506

QY 486 TGACCGTAGACTTTTCCACATGCAAGATTTTGGAACTCTTGCACTACTACTGGGAATG 545

Db 507 TGACCGTAGACTTTTCCACATGCAAGATTTTGGAACTCTTGCACTACTACTGGGAATG 566

QY 546 AGTTGCAATTAATGCTACACCATGCTGCGATCGAGTAAGTCGCTTAATGTTGTAATAA 605

Db 567 AGTTGCAATTAATGCTACACCATGCTGCGATCGAGTAAGTCGCTTAATGTTGTAATAA 626

QY 606 AGCAGAGCAAAAGGTGATGAGATGAACCTCTGTTTCATCGAATAAATAATGACTT 665

Db 627 AGCAGAGCAAAAGGTGATGAGATGAACCTCTGTTTCATCGAATAAATAATGACTT 686

RESULT 2

AW983877

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 649)

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,

Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,

Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,

O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

10737800

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=at2=RC1-HN0003-220300-021-f03&t3=2000-03-22&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 3

High quality sequence stop: 646.

location/Qualifiers

1..649

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="HN0003"

/note="Organ: head_normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 27.1%; Score 628.6; DB 10; Length 649;

Best Local Similarity 99.2%; Pred. No. 7.9e-136;

Matches 642; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 2 AGCTTTCTAATTTAACCTTTGTGAGGTACCAACTACTAAGGTTGTAGGCTCAAGAGGG 61

Db 4 AGCTTTCTAATTTAA-CTTGTGAGGTACCAACTACTAAGGTTGTAGGCTCAAGAGGG 62

QY 62 TGTGCTGTGCTAGGTAATAAAGTGAAGTGGAGCTTAATATTTCTATATTGTTGTTCT 121

Db 63 TGTGCTGTGCTAGGTAATAAAGTGAAGTGGAGCTTAATATTTCTATATTGTTGTTCT 122

QY 122 TTCTGCAAAAAGTGGGAAGTGAATGAATTTATTTCTAACAATTAATCTGCATCATA 181

Db 123 TTCTGCAAAAAGTGGGAAGTGAATGAATTTATTTCTAACAATTAATCTGCATCATA 182

QY	182	CCCTCCGAGCATTATTAAAGCAATTCGCTAATPAAGTTCTCGCTGGAAGAGTAGTTTTTTC	241
Db	183	CCCTCCGAGCATTATTAAAGCAATTCGCTAATATGTCTCGCTGGAAGAGTAGTTTTTTC	242
QY	242	AATGTACTTTAACCTTCATCTCTGTTCATTATCATCGCTTTTAAACGGTTTCGACCTTCTA	301
Db	243	AATGTACTTTATCTTCATCTCTGTTCATTATCATCGCTTTTAAACGGTTTCGACCTTCTA	302
QY	302	ATCCTATCTGACCATTAAATTTTTTAGAATGTTTCATATAGAAAGCTCTGAATCAACGG	361
Db	303	ATCCTATCTGACCATTAAATTTTTTAGAATGTTTCATATAGAAAGCTCTGAATCAACGG	362
QY	362	ACTGCGATAATAAGTGTGTGTTATCCGAATTTGTCACTTCAAGTAAACACACCTCACGAG	421
Db	363	ACTGCGATAATAAGTGTGTGTTATCCGAATTTGTCACTTCAAGTAAACACACCTCACGAG	422
QY	422	TTAAACACCTAAGTCTCACCGAATGCTCAATATCCGACGGATAATATTTATTGCTT	481
Db	423	TTAAACACCTAAGTCTCACCGAATGCTCAATATCCGACGGATAATATTTATTGCTT	482
QY	482	CTCTTGACCGTAGCACTTTCACATGCAGAGATTTTGGAAACCTCTTGCACTACTACTGGGG	541
Db	483	CTCTTGACCGTAGCACTTTCACATGCAGAGATTTTGGAAACCTCTTGCACTACTACTGGGG	542
QY	542	AATGAGTTGCAATTATTGCTACACCAATTGCGTGCAATGAGTAAGTCGCTTAATGTTGTA	601
Db	543	AATGAGTTGCAATTATTGCTACACCAATTGCGTGCAATGAGTAAGTCGCTTAATGTTGTA	602
QY	602	AAAAAGCAGAGAGCAAAAGGTGATGCAGATGAACCTCTGTTTCAATCG 648	
Db	603	AAAAAGCAGAGAGCAAAAGGTGATCCCGATGAACCTCTGTTTCAATCG 649	

RESULT 3				
AW983875				
LOCUS	AW983875	663 bp	mRNA	linear
DEFINITION	RC1-HN0003-220300-021-e02 HN0003		Homo sapiens	CDNA, mRNA sequence.

ACCESSION	AW983875
VERSION	AW983875.1
	GI:8175469

KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

REFERENCE
AUTHORS
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R.,
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 663)

Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663
 PUBMED 10737800
 COMMENT Contact: Simpson A.J.G.

Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

Tel: +55-11-2704922
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome project. This entry can be seen in the following URL
<http://www.ludwig.lu.se/>

(http://www.iudwig.org.br/scripts/gethtml2.pl?tl=<2=RC1-HN0003-2203-00-021-e02&ts=2000-03-22&ta=1)

seq primer: puc 18 forward
High quality sequence start: 15
High quality sequence stop: 661.

FEATURES

source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HN0003"
/note="Organ: head normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

```

Query Match	26.0%	Score 602.8;	DB 10;	Length 663;
Best Local Similarity	99.5%;	pred. No. 8.2e-130;		
Matches 615; Conservative	0;	Mismatches 2;	Indels 1;	Gaps 1;

QY	2	AGCTTTTCTAATTTAACCTTTGTCAAGTTACCAACTACTAAGGTTGTAAGCTCAAGAGG	61
Db	20	AGCTTTTCTAATTTAAC- TTGTCAAGTTACCAACTACTAAGGTTGTAAGCTCAAGAGG	78
QY	62	TGTGTCTGTGCTAGGTAATACTGACCTGTGAGCTTAATATTTCTATATTGTGTCT	121
Db	79	TGTGTCTGTGCTAGGTAATACTGACCTGTGAGCTTAATATTTCTATATTGTGTCT	138
QY	122	TTCTGCAAAAAAGTGGGGAAGTGAATGAATTAATTTCTAACATTTATCTGCATCATA	181
Db	139	TTCTGCAAAAAAGTGGGGAAGTGAATGAATTAATTTCTAACATTTATCTGCATCATA	198
QY	182	CCTTCCGAGCATTATTAAGCATTTGCTATTAAGTTCTCGCTGGAAGAGGTATTTTTC	241
Db	199	CCTTCCGAGCATTATTAAGCATTTGCTATTAAGTTCTCGCTGGAAGAGGTATTTTTC	258
QY	242	ATTGTACTTTACCTTCATCTCTGTTCAATTATCATCGCTTTTAAAACGGTTGACCTTCTA	301
Db	259	ATTGTACTTTACCTTCATCTCTGTTCAATTATCATCGCTTTTAAAACGGTTGACCTTCTA	318
QY	302	ATCCTATCTGACCATTATTAATTTTTAGAATGTTTCATAAGAAAGCTCTGAATCAACGG	361
Db	319	ATCCTATCTGACCATTATTAATTTTTAGAATGTTTCATAAGAAAGCTCTGAATCAACGG	378
QY	362	ACTGCGATAATAAGTGTGTGATCCAGAATTTGTCACTTCAAGTAATAAAACACCTCAGAG	421
Db	379	ACTGCGATAATAAGTGTGTGATCCAGAATTTGTCACTTCAAGTAATAAAACACCTCAGAG	438
QY	422	TTAAAAACACCTAAGTTCTACCCGAATGCTCAATATCCGGACCGATAATATTTATTTGCTT	481
Db	439	TTAAAAACACCTAAGTTCTACCCGAATGCTCAATATCCGGACCGATAATATTTATTTGCTT	498
QY	482	CTCTTGACCGTAGACTTTCCACATGCAGAGATTTTGAACCTCTTGCAGTACTACTGGGG	541
Db	499	CTCTTGACCGTAGACTTTCCACATGCAGAGATTTTGAACCTCTTGCAGTACTACTGGGG	558
QY	542	AATGAGTTGCAATTAATGCTACACCAATTGCGTGCAATCGAATGAATGCGCTTAATGTTGTA	601
Db	559	AATGAGTTGCAATTAATGCTACACCAATTGCGTGCAATCGAATGAATGCGCTTAATGTTGTA	618
QY	602	AAAAAGCAGAGAGCAAG 619	
Db	619	AAAAAGCAGAGAGCAAG 636	

RESULT 4
AW984211

LOCUS	AM984211	547 bp	mRNA	linear	EST 02-JUN-2000
DEFINITION	PM2-HN0008-120400-003-b09 HN0008 Homo sapiens cDNA, mRNA sequence.				

ACCESSION	AW9884211
VERSION	AW9884211.1
ENTRY DATE	GI:8175811

KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE	TITLE	JOURNAL	MEDLINE	PUBMED	COMMENT	FEATURES	ORIGIN
1 (bases 1 to 547)	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.						
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)	20202663	10737800	Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=at2=PM2-HN0008-120400-003-b09&t3=2000-04-12&t4=1) Seq primer: puc 18 forward High quality sequence start: 13 High quality sequence stop: 480. Location/Qualifiers 1. 547 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /dev_stage="Adult" /clone_lib="HN0008" /note="Organ: head normal; Vector: puc18; Site 1: Smal; Site 2: Smal; A mini-library was made by cloning products derived from ORESTES PCR (U.S. letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."		
1035	TCGGTCAGGTGTTCTTTAGAGAGAGTAAGAGATCAAAATGCACTAAACGAACTGAAAC						
547	TGGGTCAGCTTCTTTTGAAGAGATTAAAGGATCAATGCACTAAACGTAACGTATAC						
1095	AAGCGATCGAAATATCCCTT- GGGATTCTTGACTCGATAGTCTATTATTTTCAGAGA						
487	AAGCGATCGTAATATCCCTTTGGGGGATTCTTGACTCGATAGTCTATTATTTTCAGAGA						
1154	AAAAATATTCATTGTTTCTGGGTTGGTGATTCACCAATCATTCATTCAAATTTGTTG						
427	AAAAATATTCATTGTTTCTGGGTTGGTGATTCACCAATCATTCATTCAAATTTGTTG						
1214	TTTACCAACACCCATTCGCGCCGATAAAGCATGAATGTTGCTGGGCATAGAAATTAA						
367	TTTACCAACACCCATTCGCGCCGATAAAGCATGAATGTTGCTGGGCATAGAAATTAA						
1274	CCGTCACCTCAAAAGGTATAGTTAAATCACTGAATCCGGAGACATTTTCTATTAAATG						
307	CCGTCACCTCAAAAGGTATAGTTAAATCACTGAATCCGGAGACATTTTCTATTAAATG						
1334	AAAAGTGGAAATCTGACAATCTGGCAAAACCATTTAACACACGTCGGAAGTGTCCATGAA						
247	AAAAGTGGAAATCTGACAATCTGGCAAAACCATTTAACACACGTCGGAAGTGTCCATGAA						
1394	TTTCTGAAAGAGTTACCCCTCTAAGTAATGAGGTGTTAAGGACGCTTTCAATTTCAATGT						

DB		187	TTTCTGAAAGATTACCCCTTAAGTAATGAGGTGTTAAGAACGCCTTTCAATTTCCAATGT	128
OY		1454	CGGCTAATCGATTGGCCACTACTACTAAATCCTGAATAAGCTTTAAGAAGGTTATGTTTAA	1513
DB		127	CGGCTAATCGATTGGCCACTACTACTAAATCCTGAATAAGCTTTAAGAAGGTTATGTTTAA	68
OY		1514	AACCATCGCTTAATTTTGCTGAGATTAAACAAGTAGTCATGCTT	1557
DB		67	AACCATCGCTTAATTTTGCTGAGAGTAACATAGTAGCAATGCTT	24
RESULT 5				
BG010614				
LOCUS			446 bp	mRNA linear EST 24-JAN-2001
DEFINITION			CM4-GN0291-061200-495-d06 GN0291 Homo sapiens cDNA, mRNA sequence.	
ACCESSION			BG010614	
VERSION			BG010614.1	GI:12457996
KEYWORDS			EST.	
SOURCE			Homo sapiens (human)	
ORGANISM			Homo sapiens	
REFERENCE			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS			Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
			Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,	
			Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,	
			Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,	
			Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,	
			O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and	
			Simpson,A.U.	
TITLE			Shotgun sequencing of the human transcriptome with ORF expressed	
JOURNAL			sequence tags	
MEDLINE			Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)	
PUBMED			20202663	
COMMENT			10737800	
			Contact: Simpson A.J.G.	
			Laboratory of Cancer Genetics	
			Ludwig Institute for Cancer Research	
			Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,	
			Brazil	
			Tel: +55-11-2704922	
			Fax: +55-11-2707001	
			Email: asimpson@ludwig.org.br	
			This sequence was derived from the FAPESP/LICR Human Cancer Genome	
			Project. This entry can be seen in the following URL	
			(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM4&t2=CM4-GN0291-	
			061200-495-d06&t3=2000-12-06&t4=1)	
			Seq primer: puc 18 forward	
			High quality sequence start: 10	
			High quality sequence stop: 444.	
FEATURES			Location/Qualifiers	
source			1..446	
			/organism="Homo sapiens"	
			/mol_type="mRNA"	
			/db_xref="taxon:9606"	
			/dev_stage="Adult"	
			/clone_1lb="GN0291"	
			/note="Organ: placenta normal; Vector: puc18; Site_1:	
			Smal; Site_2: Smal; A mini-library was made by cloning	
			products derived from ORESTES PCR (U.S. Letters Patent	
			application No. 196,716 - Ludwig Institute for Cancer	
			Research) profiles into the pUC 18 vector. Reverse	
			transcription of tissue mRNA and cDNA amplification were	
			performed under low stringency conditions."	
ORIGIN				
Query Match			17.3%; Score 400.8; DB 10; Length 446;	
Best Local Similarity			96.8%; Pred. No. le-82;	
Matches 430; Conservative			0; Mismatches 12; Indels 2; Gaps 2	
OY		611	AGAGCAAAGGTGATGCAGATGAACCTCTGTTCAATCGAATAAACTAATGACTTT-CG	669
DB		4	AGCGTTAAGTGGATGCAGATGAACCTCTGG-TCATCGAATMAAACTAATGACTTTGCCG	62

QY	670	CCAAAGACATCTACTAATCTTGATAGTAATAAACAATTGCATGTCAGAGCTCATTT	729
Db	63	CCAAAGACATCTACTAATCTTGATAGTAATAAACAATTGCATGTCAGAGCTCATTT	122
QY	730	CGAAGCAGATATTTCTGATATTTGCATATAAAACAATTTAGTGAATTTATCATCGTCCACT	789
Db	123	CGAAGCAGATATTTCTGATATTTGCATATAAAACAATTTAGTGAATTTATCATCGTCCACT	182
QY	790	TGAATCTGTGTTCAATTAAGCTTAACTCTTCATATTTAGAATGAGGCTGATGAGTTCC	849
Db	183	TGAATCTGTGTTCAATTAAGCTTAACTCTTCATATTTAGAATGAGGCTGATGAGTTCC	242
QY	850	ATATTTGAAAAAGTTTTCATCACTACTAGTTTTTTGATAGCTTCAAGCCAGAAGTTGTCCTT	909
Db	243	ATATTTGAAAAAGTTTTCATCACTACTAGTTTTTTGATAGCTTCAAGCCAGAAGTTGTCCTT	302
QY	910	TTTCTATCTACTCTCATACACAACAATAAATGCTGAAATGAATCTAAGCGGAGATCGCCT	969
Db	303	TTACTATCTACTCTCATACACAACAATAAATGCTGAAATGAATCTAAGCGGAGATCGCCT	362
QY	970	AGTGATTTTAAACTATTGCTGGCAGCATTTCTGAGTCCCAATATATAAGTATTGTGTACCT	1029
Db	363	AGTGATTTTAAACTATAGCTGGCAGCATTTCTGAGTCCCAATATATAAGTATTGTGTACCT	422
QY	1030	TTTGCTGGGTCAAGGTGTTCTTTA	1053
Db	423	TTTGCTGGGTCAAGGTGTTCTTTA	446

RESULT 6	R04703/c	LOCUS	DEFINITION
	R04703	336 bp	mrna linear EST 31-MAR-1995
	pk27g02.r1	Kuwabara Mixed stage C.	briggsae Caenorhabditis briggsae
	CDNA similar to SP:VE59_IAMBD	P03754 EA59 GENE,	mrna sequence.

SOURCE ORGANISM	Caenorhabditis briggsae	Caenorhabditis briggsae
-----------------	-------------------------	-------------------------

Eukaryota; Metazoa; Chromadorea; Rhabditida; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis

REFERENCE
AUTHORS
1 (bases 1 to 336)
Hillier, L., Chissole, S., Clark, N., Couch, J.,
Dubuque, T., Holman, M., Hultman, M., Kucaba, T.,
Kuwabara, P., Le, M., Mardis, E., Marra, M., Parsons, J., Rifkin, L.,
Rohlfing, T., Tan, F., Trevaskis, E., Waterston, R., Wohldmann, P. and
Wilson, R.

TITLE	Washington University Caenorhabditis briggsae EST project
JOURNAL	Unpublished (1995)
COMMENT	Other ESTs: pk27q02.s1

Washington University Genome Sequencing Center
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1455
Fax: 314 286 1810

Email: mmarras@watson.wustl.edu
PCR_F: TGTAAACGACGGCCAGTCAGCAATTGAGCCTGG
PCR_B: CAGGAACACAGCTATGACCCCTTATGAGTATTTCTTCAGGGTA
Source: Washington University Genome Sequencing Center
PCR amplified DNA is available from Washington University Genome
Sequencing Center. Aliquots of the library may be requested from P.
Kuwabara (pek@mc-lmb.cam.ac.uk).
Seq primer: Commercially available M13 reverse dye primer.

FEATURES	Location/Qualifiers
source	1. .336

```

/organism="Caenorhabditis briggsae"
/mol_type="mRNA"
/strain="G16 Gujarati"
/db_xref="taxon:6238"
/clone_lib="Kuwabara Mixed stage C. briggsae"
/note="Vector: Lambda gt10; Site_1: EcoRI; Site_2: EcoRI;"

```

Stage:mixed, Sex:hermaphrodite. Library construction: First strand oligo(dT) primed. Second strand was as per Gubler/Hoffman. Ligated to EcoRI secondaptors. Library is non-directional. Library is non-normalized. Library constructed by P.E. Kuwabara. Additional details on construction of the library are described in P.E. Kuwabara and S. Shah, NAR 22: 4414 - 4418 (1994). Adaptor sequence: GAATTC CGTGTCTGTCG"

ORIGIN

Query Match	13.5%;	Score 314;	DB 14;	Length 336;
Best Local Similarity	99.4%;	Pred. No. 1.8e-62;		
Matches 325; Conservative	0;	Mismatches 1;	Indels 1;	Gaps 1;

QY	972	TGATTTTAAACHTATTGCTGGCAGCATTCCTTGAGTCCAATATATAAAGTATTTGTACCTTT	1031
Db	336	TGATTTTAAACTATTTGCTGGCAGCATTCCTTGAGTCCAATATATAAAGTATTTGTACCTTT	277
QY	1032	TGCTGGGTCAGGTTGTTCTTTAGGAGAGTAAAGATCAAAATGCACCTAAACGAACTGA	1092
Db	276	TGCTGGGTCAGGTTGTTCTTTAGGAGAGTAAAGATCAAAATGCACCTAAACGAACTGA	217
QY	1092	AACAAGCGATCGAAAAATATCCCTTTGGATTCTTGACTCGATAAGTCTATTATTTTCAGA	1151
Db	216	AACAAGCGATCGAAAAATAT - CCTTTGGATTCTTGACTCGATAAGTCTATTATTTTCAGA	158
QY	1152	GAAAAAATATTCATTGTTTTCTGGTGTGTGATTCACCAATCATTCATTCACAAATTGT	1211
Db	157	GAAAAAATATTCATTGTTTTCTGGTGTGTGATTCACCAATCATTCATTCACAAATTGT	98
QY	1212	TGTTTTCACCAACCCATTCCGCCGATAAAGCATGAATGTTGCTGGGCATAGAAATT	1271
Db	97	TGTTTTCACCAACCCATTCCGCCGATAAAGCATGAATGTTGCTGGGCATAGAAATT	38
QY	1272	AACCGTCACTCAAAAAGGTATAGTTAA	1298
Db	37	AACCGTCACTCAAAAAGGTATAGTTAA	11

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
BF737769	BF737769	CM4-KT0035-131200-610-a12	BF737769.1	EST.	Homo sapiens	Homo sapiens
	365 bp	KT0035	GI:12064445			
	mRNA					
	linear					
		Homo sapiens				
		CDNA,				
	EST 10-JAN-2001					
	mRNA sequence.					

REFERENCE
1 (bases 1 to 365)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominae; Homo.

AUTHORS
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE	20202663
PUBMED	10737800
COMMENT	Contact: Simpson A.J.G.

COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM4&ct=CM4-KT0035->

131200-610-a12&t3=2000-12-13&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 18
High quality sequence stop: 364.
Location/Qualifiers

1. .365
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_1lb="KT0035"

/note="Organ: bladder tumor; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 13.3%; Score 309.6; DB 10; Length 365;
Best Local Similarity 98.7%; Pred. No. 1.9e-61;
Matches 312; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 653 AAACATAAGACTTTTCGCCAAGACATCTACTAATCTTGATAGTAATAATAACAATTG 712
Db 50 AAACATAAGACTTTTCGCCAAGACATCTACTAATCTTGATAGTAATAATAACAATTG 109
QY 713 CATGTCCAGAGCTCATTCGACAGACATATTTCTGATATTGTCATATAAACAATTAGTGA 772
Db 110 CATGTCCAGAGCTCATTCGACAGACATATTTCTGATATTGTCATATAAACAATTAGTGA 169
QY 773 ATTTATCATCGTCCACTTGAATCTGTGGTTCATACGCTTAACTCTTCATATTTAGAAA 832
Db 170 ATTTATCATCGTCCACTTGAATCTGTGGTTCATACGCTTAACTCTTCATATTTAGAAA 229
QY 833 TGAGGCTGATGAGTTCATATTTGAAAAAGTTTTCATCAGTACTAGTTTGTGATAGCTT 892
Db 230 TGAGGCTGATGAGTTCATATTTGAAAAAGTTTTCATCAGTACTAGTTTGTGATAGCTT 289
QY 893 CAAGCCAGAGTGTCTTTTCTAATCTAATCTCATACCAACAATAATGCTGAATGAAAT 952
Db 290 CAAGCCAGAGTGTCTTTTCTAATCTAATCTCATACCAACAATAATGCTGAATGAAAT 349
QY 953 CTAGCGGAGATCGCC 968
Db 350 CTAGCGGAGATCGCC 365

RESULT 8
AM984213/c 262 bp mRNA linear EST 02-JUN-2000
LOCUS PM2-HN0008-120400-003-d07 HN0008 Homo sapiens cDNA, mRNA sequence.
DEFINITION
ACCESSION AM984213
VERSION AM984213.1 GI:8175813
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
DEFINITION Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 262)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.

Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=PM2-HN0008-120400-003-d07&t3=2000-04-12&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 13
High quality sequence stop: 218.
Location/Qualifiers

FEATURES

source

1. .262
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_1lb="HN0008"
/note="Organ: head normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 10.3%; Score 238.6; DB 10; Length 262;
Best Local Similarity 98.1%; Pred. No. 6.7e-45;
Matches 252; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1311 GGGAGCACTTTTCTATTAATGAAAAGTGAAATCTGCAATTTCTGGCAAAACCATTTAA 1370
Db 262 GGGAGCACTTTTCTATTAATGAAAAGTGAAATCTGCAATTTCTGGCAAAACCATTTAA 203
QY 1371 CACACGTGCGAAGTGTCCATGAATTTCTGAAAAGATTACCCCTCTAAGTAATGAGGTGTT 1430
Db 202 CACACGTGCGAAGTGTCCATGAATTTCTGAAAAGATTACCCCTCTAAGTAATGAGGTGTT 143
QY 1431 AAGGACGCTTTCATTTTCAATGTCGGCTAATCGATTGGCCATACTACTAATCCTGAAT 1490
Db 142 AAGGACGCTTTCATTTTCAATGTCGGCTAATCGATTGGCCATACTACTAATCCTGAAT 83
QY 1491 AGCTTTAAGAGGTTATGTTTAAACCATCGCTTAATTTGCTGAGATTACATAGTAGTC 1550
Db 82 AGCTTTAAGAGGTTATGTTTAAACCATCGCTTAATTTGCTGAGATTACATAGTAGTC 24
QY 1551 AATGCTTTCACCTAAGG 1567
Db 23 AATGCTTTCACATTCGG 7

RESULT 9
BF773183 254 bp mRNA linear EST 12-JAN-2001
LOCUS CM4-IT0046-151200-606-b11 IT0046 Homo sapiens cDNA, mRNA sequence.
DEFINITION
ACCESSION BF773183
VERSION BF773183.1 GI:12121083
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
DEFINITION Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 254)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

TITLE
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE
20202663

PUBMED
10737800

COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM4&t2=CM4-IT0046-
151200-606-b1&t3=2000-12-15&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 7
High quality sequence stop: 253.
Location/Qualifiers

FEATURES
Source
1. 254
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="IT0046"
/note="Organ: epid tumor; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN
Query Match 8.9%; Score 205.8; DB 10; Length 254;
Best Local Similarity 98.0%; Pred. No. 3e-37;
Matches 240; Conservative 0; Mismatches 2; Indels 3; Gaps 3;

QY 617 AAGGTGATGCAGATGAACCTCTGTTTCATCGAATAAAGTAATGACTTTCCGCAAGA 676
|||
Db 11 AAGGTGATGCAGATGAACCTTTGG-TCATCGAATCAACTAATGACTTTCCGCAAGA 69
677 CATCTACTAATCT-TGTGATAGTAATAAACAATTCGATGTCAGAGCTCATTCGAAGC 735
|||
Db 70 CATCTACTAATCTCTGTGATAGTAATAAACAATTCGATGTCAGAGCTCATTCGAAGC 129
736 AGATATTTCTGGATATGT-CATTAACAATTTAGTGAATTTATCATGTCACCTGAAT 794
|||
Db 130 AGATATTTCTGGATATGTTCATTAACAATTTAGTGAATTTATCATGTCACCTGAAT 189
795 CTGTGTTTCATTACGTCCTTAACCTCTCATATTTAGAAATGAGGCTGATGATTCATATT 854
|||
Db 190 CTGTGTTTCATTACGTCCTTAACCTCTCATATTTAGAAATGAGGCTGATGATTCATATT 249
QY 855 TGAAG 859
|||
Db 250 TGAAG 254

RESULT 10
LOCUS BF348952 206 bp mRNA linear EST 22-NOV-2000
DEFINITION MR1-DT0059-110700-002-g08 DT0059 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF348952
VERSION BF348952.1 GI:11308026
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 206)

AUTHORS
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

TITLE
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE
20202663

PUBMED
10737800

COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR1&t2=MR1-DT0059-
110700-002-g08&t3=2000-07-11&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 206.
Location/Qualifiers

FEATURES
Source
1. 206
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="DT0059"
/note="Organ: denis drash; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN
Query Match 8.7%; Score 202.8; DB 10; Length 206;
Best Local Similarity 99.0%; Pred. No. 1.5e-36;
Matches 204; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1410 CCCTTAAGTAATGAGGTGTTAAGACGCTTTCATTTCAATGTCGGCTAATCGATTGG 1469
|||
Db 1 CCCTTAAGTAATGAGGTGTTAAGACGCTTTCATTTCAATGTCGGCTAATCGATTGG 60
1470 CCATACTACTAATCCTGAATAGCTTTAAGAGGTTATGTTAAACCATCGCTTAATTT 1529
|||
Db 61 CCATACTACTAATCCTGAATAGCTTTAAGAGGTTATGTTAAACCATCGCTTAATTT 120
1530 GCTGAGATTACATAGTAGTCATGCTTTCACCTAAGAGAAAAACATTTCAGGAGTTG 1589
|||
Db 121 GCTGAGATTACATAGTAGTCATGCTTTCACCTAAGAGAAAAACATTTCAGGAGTTG 180
QY 1590 ACTGAATTTTATCTAATTAATGAAT 1615
|||
Db 181 ACTGAATTTTATCTAATTAATGAAT 206

RESULT 11
LOCUS CNS0039G/c 1101 bp DNA linear GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL063921
VERSION AL063921.1 GI:4941778
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster

Db	671	WTHAYHWATYHNYUYMYCAMMCMSTHTCHNCYUUYHNYHTAHNTHTHHMYAHYUWUYMMA	730
QY	1777	GGATATACCCGCC-----TCTTCAATACASTAACTCCAATATAGTAACCT	1828
Db	731	YUWMYCTACTYHYNHNNHNYHWAUHTTWYAWAHAMMMHNAHAAADAAWAATHTHYNHT	790
QY	1829	TAATTTTATTAADAATAACCGCAATTATTTGGCGGACACAGATCTCTTTAACTT	1888
Db	791	THYMHNTYMYHYWYTCCSYMCTYHCHWYHNTAYTCTWTNNHMYTWTWYHNTHTWNNHTT	850
QY	1889	ACTCTATATACATACAGSTTTTCCATCTAAADATYAGTAGTGAACSTTAACGGGAC	1948
Db	851	HWAMHNTHTWCWMMWHAATTWTWATGHCWACMTMHNHNNHNNHMACHAHNTHTMCSCH	910
QY	1949	GTATGTAGTTTTCSATATTTAGCTTTCTGCTTCTTTGGATAACCCAGTGTATTCAT	2008
Db	911	HMHCTCHNNHTMYHNTCHMWMHNMHNMHNMWMAWTWTMTTMMHCCMMHNNCHMNMH	970
QY	2009	GTTGCATGGTGCACSTGTTTATACCAAGCATATAGTCTATTAATGCATATATAGTATGCC	2068
Db	971	MYMYCCHYUCTCHTATHTHNYHMYCTCYHUCTHWTYTAWYMAWTAHAMTTATTWMMHMY	1030
QY	2069	GAACGATTAGCTTTCAAGGCTTCTGAAGAAGCGTTTCAAGTACTATTA	2117
Db	1031	AHWATWMMWMMWMAATAWMACTCHNHTWYHNTCTWYHNTYHNMWAMWMA	1079

RESULT 13	LOCUS	DEFINITION
BB755979	BB755979	556 bp mRNA linear
BB755979	RIKEN full-length enriched, melanocyte clone G270069113 3', mRNA sequence.	EST 17-OCT-2001 musculus CDNA

ACCESSION	BB755979	
VERSION	BB755979.1	GI:16202468
KEYWORDS	EST.	
SOURCE	Mus musculus	(house mouse)
ORGANISM	Mus musculus	

REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
1 (bases 1 to 556)
Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanaqaki, T.,

Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Imotani, K., Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Wataniki, A., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al. 2001)

JOURNAL
COMMENT

Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>,
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagti, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,
Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.
and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)

Komno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.
e mouse tissues.

FEATURES
source

```

1. .556
/organism="Mus musculus"
/mol_type="mrna"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="G270069I13"
/cell_type="melanocyte"
/clone_lib="RIKEN full-length enriched, melanocyte"

```

ORIGIN

Query Match	2.88;	Score 65.4;	DB 10;	Length 556;
Best Local Similarity	98.5%;	Pred. No. 0.00017;		
Matches 66;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;

Oy 1 AAGCTTTCTAATTAAACCTTGTGAGGTTAACCACTACTAAGCTTGTAAGCTCAGAAGG 60
|||
Db 490 AAGCTTTCTAATTAAACCTTGTGAGGTTAACCACTACTAAGCTTGTAAGCTCAGAAG 549

QY	61	GTGTGTC	67
Db	550	GTGTGTC	556

RESULT 14	
BX414650	
LOCUS	BX414650
DEFINITION	Homo sapiens THYMUS Homologous to human thymus gland cDNA clone CS0CAP001YN02
	994 bp mRNA linear EST 15-MAY-2003
	3-PRIME, mRNA sequence.

ACCESSION	BX414650	GI:30763455
VERSION	BX414650.1	
KEYWORDS	EST.	
SOURCE	Homo sapiens	(human)
ORGANISM	Homo sapiens	

REFERENCE	1 (bases 1 to 994)
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	Contact: Genoscope

Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: secref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6015.f
Contact : Feng Liang Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com/InvitrogenCorporation> 16000
Faraday Avenue Genoscope sequence ID : CS0CAP001DG0INPL.

FEATURES

```

1. .994
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0CAP001YN02"
/tissue_type="THYMUS"
/clone_lib="Homo sapiens THYMUS"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched, into
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

```

ORIGIN

Query Match 2.6%; Score 61; DB 13; Length 994;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 6, 2004, 05:00:55 ; Search time 900 Seconds
(without alignments)
10960.351 Million cell updates/sec

Title: US-10-074-744-4
Perfect score: 2322
Sequence: 1 aagcttttctaatttaacct.....gtcgtagtgggtgcacaca 2322

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn19808:*
2: geneseqn19908:*
3: geneseqn20008:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2322	100.0	2322	7	ABX15851	Abx15851 Bacteriop
2	2319.4	99.9	9980	7	ACA55352	Acas55352 PIAO-P/L
3	661.2	28.5	1469	7	ABQ80306	Abq80306 Lambda fr
4	661.2	28.5	1469	7	ACC83020	Acc83020 Bacteriop
5	75	3.2	10523	8	ACD19000	Acid19000 E. coli 0
6	75	3.2	23480	9	ADC01049	Adc01049 Enteroha
7	71.4	3.1	9393	6	ABS78855	Abs78855 E. coli C
8	50.8	2.2	2000	7	ADA71938	Ada71938 Rice gene
9	49.8	2.1	2000	7	ADA71938	Ada71938 Rice gene
10	45.4	2.0	50000	6	ABU56201	Abu56201 Armpv gen
11	44.2	1.9	453	6	ABU79740	Abu79740 Human ova
12	44.2	1.9	5739	4	AAK78745	Aak78745 Human imm
13	44.2	1.9	7416	4	AAK78744	Aak78744 Human imm
14	44.2	1.9	33513	4	AAK78746	Aak78746 Human imm
15	43.6	1.9	110000	6	ABQ69245_14	Abq69245_14
16	43.6	1.9	110000	6	ABQ67195_2	Abq67195_2
17	43.4	1.9	1572	7	ACA23191	Aca23191 Prokaryot
18	43.4	1.9	110000	2	AAK20248_01	AAK20248_01
19	42.8	1.8	955	3	AAZ33364	Aaz33364 Human sec
20	42.8	1.8	13202	6	ABU333485	Abu333485 Human imm
21	42.2	1.8	6308	6	ABU33470	Abu33470 Human imm
22	42	1.8	110000	2	AAV21209_05	AAV21209_05
23	41.6	1.8	13784	6	ABK40061	Abk40061 Human che

C	24	41.4	1.8	6131	6	ABL32891	Ab132891 Human imm
	25	41.4	1.8	110000	5	ABQ74964_0	Abq74964 Human kin
	26	41.2	1.8	458	5	ABV13735	Abv13735 Human pro
	27	41.2	1.8	506	5	ABV04566	Abv04566 Human pro
	28	41.2	1.8	533	5	ABV43698	Abv43698 Human pro
	29	41.2	1.8	533	5	ABV34848	Abv34848 Human pro
	30	41.2	1.8	6907	6	ABL32567	Ab132567 Human imm
C	31	41	1.8	5328	6	ABL32561	Ab132561 Human imm
C	32	41	1.8	6309	6	ABL32304	Ab132304 Human imm
C	33	41	1.8	110000	7	ACF67367_47	Continuation (48 O
	34	41	1.8	110000	7	ACF65388	Acf65388 Phototrab
C	35	40.8	1.8	711	7	ACA52657	Acas52657 Prokaryot
C	36	40.8	1.8	6154	6	ABK31201	Abk31201 Signal tr
C	37	40.8	1.8	6154	6	ABL70168	Ab170168 Chemicall
C	38	40.8	1.8	6154	6	AA561113	Aas61113 Human gen
	39	40.8	1.8	110000	2	AAK20248_00	Aax20248 Borrelia
C	40	40.8	1.8	113515	6	ABL34175	Ab134175 Human imm
C	41	40.6	1.7	5994	6	ABL32879	Ab132879 Human imm
C	42	40.6	1.7	6078	6	ABK28228	Abk28228 DNA trans
C	43	40.4	1.7	3926	4	AAH54936	Aah54936 S. epider
C	44	40.4	1.7	4590	5	AAH24065	Aah24065 Yeast AOD
	45	40.4	1.7	5775	6	ABQ70927	Abq70927 Listeria

ALIGNMENTS

RESULT 1
ID ABX15851
ABX15851 standard; DNA; 2322 BP.
XX
AC ABX15851;
XX
DT 10-APR-2003 (first entry)
XX
DE Bacteriophage lambda spacer polynucleotide #3.
XX
KW Genetic engineering; ds; transcriptional interference; genomics;
KW gene expression cassette; plant; herbicide resistance; nutrient content;
KW fungicide resistance; insect resistance; drought-tolerance;
KW spacer polynucleotide.
XX
OS Bacteriophage lambda.
XX
PN US2002155540-A1.
XX
XX
PD 24-OCT-2002.
XX
PF 13-FEB-2002; 2002US-00074744.
XX
PR 14-FEB-2001; 2001US-0268584P.
XX
PA (PADI/) PADIDAM M.
XX
XX Padidam M;
XX
DR WPI; 2003-198498/19.
XX
PT Reducing transcriptional interference between tandemly arranged gene
PT expression cassettes in a cell, by modifying orientation of the cassettes
PT and/or by placing a spacer polynucleotide sequence between the cassettes.
XX
PS Claim 3; Page 18-19; 33pp; English.
XX
CC This invention relates to a novel method for reducing transcriptional
CC interference (TI) between two or more tandemly arranged gene expression
CC cassettes in a host cell, comprising introducing into the cell a
CC polynucleotide comprising a gene expression cassette encoding a first
CC polypeptide, a spacer polynucleotide and a gene expression cassette
CC encoding a second polypeptide, and culturing the cell, where
CC transcriptional interference between the two gene expression cassettes is
CC reduced and polypeptides are expressed. The method of the invention is
CC useful for reducing or eliminating transcriptional interference between

CC adjacent genes in a host such as bacteria, fungus, yeast, animal, mammal
CC or plant cell (e.g. apple, Arabidopsis, bajra, banana, barley, bean,
CC beet, blackgram, chickpea, chili, cucumber, eggplant, fabbean, maize,
CC melon, millet, oat, akra, papaya, peanut, pea, pepper, pigeonpea,
CC pineapple, pumpkin, rice, squash, sugarcane, sunflower, sweet potato,
CC tea, tomato, tobacco, watermelon or wheat). to improve expression of
CC multiple gene expression constructs in both prokaryotic and eukaryotic
CC host cells. Gene expression in transgenic plant host cells or transgenic
CC plants is useful for producing various desirable traits or gene products,
CC including control of flowering, herbicide resistance, fungicide
CC resistance, insect resistance, plant size or form, nutrient content,
CC drought-tolerance, various pathway intermediates for the modulation of
CC pathways already existing in the host for the synthesis of new products,
CC and examination and determination of ectopic gene expression and gene
CC function, particularly as related to the fields of genomics and
CC proteomics. The present sequence represents a bacteriophage lambda spacer
CC polynucleotide used in the method of the invention

XX Sequence 2322 BP; 694 A; 446 C; 415 G; 767 T; 0 U; 0 Other;

Query Match 100.0%; Score 2322; DB 7; Length 2322;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGCTTTTCTAATTAAACCTTTGTCAGGTTACCACTACTAAGGTTGAGGCTCAAGAGG 60
Db 1 AAGCTTTTCTAATTAACTTTGTCAGGTTACCACTACTAAGGTTGAGGCTCAAGAGG 60
QY 61 GTGTGTCCTGTCGTAGGTAATAAATGACCTGTGAGCTTAATAATTCTATATTGTTGTC 120
Db 61 GTGTGTCCTGTCGTAGGTAATAAATGACCTGTGAGCTTAATAATTCTATATTGTTGTC 120
QY 121 TTTCGCAAAAAAGTGGGGAAGTGAATGAATAATTCTTAACATTTATCTGCATCAT 180
Db 121 TTTCGCAAAAAAGTGGGGAAGTGAATGAATAATTCTTAACATTTATCTGCATCAT 180
QY 181 ACCTTCCGAGCATTTATTAAGCATTTCCGCTATAAGTTCGCGTGAAGAGGAGTAGTTTTT 240
Db 181 ACCTTCCGAGCATTTATTAAGCATTTCCGCTATAAGTTCGCGTGAAGAGGAGTAGTTTTT 240
QY 241 CATTTGACTTTAACCCTTCATCTCTGTTCAATATCATCGCTTTTAAACGGTTGACCTTCT 300
Db 241 CATTTGACTTTAACCCTTCATCTCTGTTCAATATCATCGCTTTTAAACGGTTGACCTTCT 300
QY 301 AATCCTATCTGACCATTAATAATTTTTAGATGTTTCATAGAAGAGCTCTGAATCAAG 360
Db 301 AATCCTATCTGACCATTAATAATTTTTAGATGTTTCATAGAAGAGAGCTCTGAATCAAG 360
QY 361 GACTGCGATAATAGTGTGTGATCCAGAATTTGTCACTTCAAGTAAACACCTCAGCA 420
Db 361 GACTGCGATAATAGTGTGTGATCCAGAATTTGTCACTTCAAGTAAACACCTCAGCA 420
QY 421 GTTAAACACCTAAGTTCACCGAATGTTCTCAATATCCGACGGATAATATTATTATGCT 480
Db 421 GTTAAACACCTAAGTTCACCGAATGTTCTCAATATCCGACGGATAATATTATTATGCT 480
QY 481 TCTCTTGACCGTAGGACTTTCACATGCAAGATTTTGAACCTCTTGCACTACTACTGGG 540
Db 481 TCTCTTGACCGTAGGACTTTCACATGCAAGATTTTGAACCTCTTGCACTACTACTGGG 540
QY 541 GAATGAGTTGCAATTATTTGCTACACCATTTGCGTGCATCGAGTAAAGTGGCTTAATGTTGCT 600
Db 541 GAATGAGTTGCAATTATTTGCTACACCATTTGCGTGCATCGAGTAAAGTGGCTTAATGTTGCT 600
QY 601 AAAAAAGCAGAGCAAAAGTGGATGCAATGAACCTCTGTTTCATGCAATAAAACTAAT 660
Db 601 AAAAAAGCAGAGCAAAAGTGGATGCAATGAACCTCTGTTTCATGCAATAAAACTAAT 660
QY 661 GACTTTTCGCCAAGACATCTAATCTTGTGATAGTAAATAAACAATTGCATGTCCA 720
Db 661 GACTTTTCGCCAAGACATCTAATCTTGTGATAGTAAATAAACAATTGCATGTCCA 720
QY 721 GAGCTCATTTGGAAGCAGATATTTCTGATATGTCAATAAACAAATTAGTAATTATCA 780

Db 721 GAGCTCATTTGGAAGCAGATATTTCTGATATGTCAATAAACAAATTAGTAATTATCA 780
QY 781 TCGTCCACTTGAATCTGTGTTTCATTAACGTTCTTAACCTCTTCATATTTAGAATGAGGCTG 840
Db 781 TCGTCCACTTGAATCTGTGTTTCATTAACGTTCTTAACCTCTTCATATTTAGAATGAGGCTG 840
QY 841 ATGAGTTCATATTTGAAAAGTTTTCATCACTACTTATTTTTGATAGCTCAAGCCAG 900
Db 841 ATGAGTTCATATTTGAAAAGTTTTCATCACTACTTATTTTTGATAGCTCAAGCCAG 900
QY 901 AGTTGCTTTTCTATCTACTCTCATACCAACAATAAATGCTGAATGAATTCTAAGCGG 960
Db 901 AGTTGCTTTTCTATCTACTCTCATACCAACAATAAATGCTGAATGAATTCTAAGCGG 960
QY 961 AGATCGCCTAGTGAATTTAACTAATTGCTGCGAGCATTTGAGTCCAAATATATAAGTAT 1020
Db 961 AGATCGCCTAGTGAATTTAACTAATTGCTGCGAGCATTTGAGTCCAAATATATAAGTAT 1020
QY 1021 TGTGTACCTTTTGTGCTGGGTGAGGTTGTTCTTTAGAGAGGTAAAGATCAAAATGCACCTA 1080
Db 1021 TGTGTACCTTTTGTGCTGGGTGAGGTTGTTCTTTAGAGAGGTAAAGATCAAAATGCACCTA 1080
QY 1081 AAGCAAACTGAAACAAGCGATCGAAAATATCCCTTGGGATTTCTGACTCGATAAGTCTA 1140
Db 1081 AAGCAAACTGAAACAAGCGATCGAAAATATCCCTTGGGATTTCTGACTCGATAAGTCTA 1140
QY 1141 TTATTTTCAGAGAAAAAATATTCATTTGTTTCTGCGTGGTGTGATGCAACCAATCATTCGA 1200
Db 1141 TTATTTTCAGAGAAAAAATATTCATTTGTTTCTGCGTGGTGTGATGCAACCAATCATTCGA 1200
QY 1201 TTCAAAATTTGTTGTTTACCAACCCATTCCGCCCGATAAAAAGCATGAATGTTGCTGCTG 1260
Db 1201 TTCAAAATTTGTTGTTTACCAACCCATTCCGCCCGATAAAAAGCATGAATGTTGCTGCTG 1260
QY 1261 GGCATAGAATTAACCGCTACCTCAAAAAGGTATAGTTAAATCACTGAATCCGGAGACATT 1320
Db 1261 GGCATAGAATTAACCGCTACCTCAAAAAGGTATAGTTAAATCACTGAATCCGGAGACATT 1320
QY 1321 TTTCTATTAATGAAAAAGTGAATACTGACAATTTGCGCAACCATTTAACACACGTCGG 1380
Db 1321 TTTCTATTAATGAAAAAGTGAATACTGACAATTTGCGCAACCATTTAACACACGTCGG 1380
QY 1381 AACTGTCCATGAATTTCTGAAAAGATTACCCCTCTAAGTATGAGTGTTAAGACGCTT 1440
Db 1381 AACTGTCCATGAATTTCTGAAAAGATTACCCCTCTAAGTATGAGTGTTAAGACGCTT 1440
QY 1441 TCATTTTCAATGTCGGCTAATCGATTTGGCCATACCTAATAATCCTGAATAGCTTTAGA 1500
Db 1441 TCATTTTCAATGTCGGCTAATCGATTTGGCCATACCTAATAATCCTGAATAGCTTTAGA 1500
QY 1501 AGGTTATGTTTAAACCATGCTTAATTTGCTGAGATTAACATAGTAGTCAATGCTTTCA 1560
Db 1501 AGGTTATGTTTAAACCATGCTTAATTTGCTGAGATTAACATAGTAGTCAATGCTTTCA 1560
QY 1561 CCTAAGAAAAAAACATTTCAAGGAGTTGACTGAATTTTATCTAATTAATGAATAAGTG 1620
Db 1561 CCTAAGAAAAAAACATTTCAAGGAGTTGACTGAATTTTATCTAATTAATGAATAAGTG 1620
QY 1621 CTTACTTCTTCTTTTGAACCTAACAACCAATTTTAACTTCCGATATCGCAATTTTCA 1680
Db 1621 CTTACTTCTTCTTTTGAACCTAACAACCAATTTTAACTTCCGATATCGCAATTTTCA 1680
QY 1681 CCATGCTCATCAAAAGACAGTAAGATAAAACATTTGTAACAAAGATAGTCATTCCAAACA 1740
Db 1681 CCATGCTCATCAAAAGACAGTAAGATAAAACATTTGTAACAAAGATAGTCATTCCAAACA 1740
QY 1741 TCTGCTCGTAGGAATGCCCTAATTTTTTCTACTGCAAGAAATATACCCGCCCTCTTCAATA 1800
Db 1741 TCTGCTCGTAGGAATGCCCTAATTTTTTCTACTGCAAGAAATATACCCGCCCTCTTCAATA 1800
QY 1801 ACACTAAACTCCAACATATAGTAACCCCTTAATTTTATTAATAACCGCAATTATTTGG 1860

Db 1801 AACTAACTCCACATATAGTAACCTTAATTTTATTAATAAACCGCAATTATTTGG 1860
QY 1861 CGGCAACACAGGATCTCTTTTAACTTCTCTATATACATACGTTTCCATCTAATAA 1920
Db 1861 CGGCAACACAGGATCTCTTTTAACTTCTCTATATACATACGTTTCCATCTAATAA 1920
QY 1921 TTAGTAGTATGAAGTAAACGGGGGATCGTATTGTAGTTTCCATATTAGCTTCTGCT 1980
Db 1921 TTAGTAGTATGAAGTAAACGGGGGATCGTATTGTAGTTTCCATATTAGCTTCTGCT 1980
QY 1981 TCCCTTTGGATTAACCCACTGTATTCATGTTGCATGTTGCACGTTTATACCAACGATAT 2040
Db 1981 TCCCTTTGGATTAACCCACTGTATTCATGTTGCATGTTGCACGTTTATACCAACGATAT 2040
QY 2041 AGTCTAATTAATGATATATAGTATCGCCGAACGATTAGCTCTTCAAGGCTTGAAGAAC 2100
Db 2041 AGTCTAATTAATGATATATAGTATCGCCGAACGATTAGCTCTTCAAGGCTTGAAGAAC 2100
QY 2101 GTTTCAGTACTAATAAGCCGATAGATAGCCACGACCTTCGTAGCCATTTTTCATAAGTG 2160
Db 2101 GTTTCAGTACTAATAAGCCGATAGATAGCCACGACCTTCGTAGCCATTTTTCATAAGTG 2160
QY 2161 TTAACCTCCGCTCCTCGCTCATTAACAGACATTCACCTACATGTTATGGCGAAAGTATGCA 2220
Db 2161 TTAACCTCCGCTCCTCGCTCATTAACAGACATTCACCTACATGTTATGGCGAAAGTATGCA 2220
QY 2221 TCGTGGGTGTGGGAAAGTCGTGAAGAAAGAAAGAGTCAAGCTGCTGTTGACATCAGTGC 2280
Db 2221 TCGTGGGTGTGGGAAAGTCGTGAAGAAAGAAAGAGTCAAGCTGCTGTTGACATCAGTGC 2280
QY 2281 TATCTTCTTACTGTTATGACGTCGTAGTGGGTGGCACACA 2322
Db 2281 TATCTTCTTACTGTTATGACGTCGTAGTGGGTGGCACACA 2322
RESULT 2
ACA55352
ID ACA55352 standard; DNA; 9980 BP.
XX ACA55352;
AC ACA55352;
XX
DT 06-JUN-2003 (first entry)
XX
DE PIAO-P/L plasmid series, pIAO-p/L-Lambda nucleotide sequence.
XX
KW PiggyBac; transposon; eukaryotic transformation vector; ds;
KW transformed cell; transformed embryo; transgenic; plasmid; cyclic;
KW circular.
XX
OS Synthetic.
XX
PN US2002173634-A1.
XX
PD 21-NOV-2002.
XX
PF 30-OCT-2001; 2001US-00001189.
XX
PR 31-OCT-2000; 2000US-0244677P.
PR 01-NOV-2000; 2000US-0244984P.
XX
PA (FRAS/) FRASER M J.
PA (LIXX/) LI X.
PA (BEAM/) BEAM T.
PA (HUA-VAN/) HUA-VAN A.
XX
PI Fraser MJ, Li X, Beam T, Hua-Van A;
XX
DR WPI; 2003-352597/33.
DR P-PSDB; ABU70356.
XX
PT New DNA molecule in the transposon piggyBac, useful for transferring
PT genes into host cells or embryos for transforming the cells of embryos
PT that can be used in making transgenic organisms.

XX Example 2; Fig 2(C2); 151pp; English.
PS
XX The invention describes a DNA molecule comprising at least 163
CC consecutive nucleotide base pairs of the 3' terminal region beginning at
CC the 3' terminal base pair, and at least 125 consecutive nucleotide base
CC pairs of the 5' terminal region beginning at the 5' terminal base pair, of
CC the piggyBac molecule. The region extends from the restriction site SacI
CC to the end of the piggyBac molecule. The DNA molecule in the transposon
CC piggyBac is useful for transferring genes into host cells or embryos for
CC transforming the cells of embryos. The transformed cells or embryos are
CC useful for developing or making transgenic organisms. This sequence
CC represents plasmid used in the creation of minimal sequence eukaryotic
CC transformation vector piggyBac
XX
SQ Sequence 9980 BP; 2697 A; 2342 C; 2139 G; 2801 T; 0 U; 1 Other;

Query Match 99.9%; Score 2319.4; DB 7; Length 9980;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2320; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGCTTTCTAATTTAACTTGTCAAGTTACCACTAAGTGTAGGCTCAAGAGG 61
Db 6608 AGCTTTCTAATTTAACTTGTCAAGTTACCACTAAGTGTAGGCTCAAGAGG 6667
QY 62 TGTGTCCCTGCTAGTAATAAAGTACCTGTGAGCTTAATATCTATATGTTGCT 121
Db 6668 TGTGTCCCTGCTAGTAATAAAGTACCTGTGAGCTTAATATCTATATGTTGCT 6727
QY 122 TTCTGCAAAAAAGTGGGAGTGAATGAATTAATTTCTAATTTATCTGCATCAT 181
Db 6728 TTCTGCAAAAAAGTGGGAGTGAATGAATTAATTTCTAATTTATCTGCATCAT 6787
QY 182 CCTTCCGAGCATTTATTAAAGCATTTGCTATTAAGTTCTGCTGGAAGAGTATTTTC 241
Db 6788 CCTTCCGAGCATTTATTAAAGCATTTGCTATTAAGTTCTGCTGGAAGAGTATTTTC 6847
QY 242 ATTGACTTTACCTTCACTCTGCTGCTATTCATTCATGCTTTTAAACGGTTGACCTTCT 301
Db 6848 ATTGACTTTACCTTCACTCTGCTGCTATTCATTCATGCTTTTAAACGGTTGACCTTCT 6907
QY 302 ATCTATCTGACCATTAATTTTAAAGTGTTCATTAAGAAAGCTCTGAATCAAGG 361
Db 6908 ATCTATCTGACCATTAATTTTAAAGTGTTCATTAAGAAAGCTCTGAATCAAGG 6967
QY 362 ACTGGAATATAAGTGTGTATCCAGAAATTTGTCACTTCAAGTAAACACCTCAGAG 421
Db 6968 ACTGGAATATAAGTGTGTATCCAGAAATTTGTCACTTCAAGTAAACACCTCAGAG 7027
QY 422 TTAACAACCTAAGTCTCACCGAAATGTCTCAATATCCGACGGAATATTTATTTGCT 481
Db 7028 TTAACAACCTAAGTCTCACCGAAATGTCTCAATATCCGACGGAATATTTATTTGCT 7087
QY 482 CTCTGACCGTAGGACTTTCCACATGACGATTTTGAACCTCTTGACGACTACTGGGG 541
Db 7088 CTCTGACCGTAGGACTTTCCACATGACGATTTTGAACCTCTTGACGACTACTGGGG 7147
QY 542 AATGAGTTGCAATTATTGCTACACCATTTGCGTCATCGAGTAAGTGGCTTAATGTTGTA 601
Db 7148 AATGAGTTGCAATTATTGCTACACCATTTGCGTCATCGAGTAAGTGGCTTAATGTTGTA 7207
QY 602 AAAAAGCAGAGCAAAAGTGGATGACATGAACCTCTGTTCAATGAATAAACTAATG 661
Db 7208 AAAAAGCAGAGCAAAAGTGGATGACATGAACCTCTGTTCAATGAATAAACTAATG 7267
QY 662 ACTTTTCCCAACGACATCTAATATCTTGTGATAGTAATAAACAATTCATGTCAG 721
Db 7268 ACTTTTCCCAACGACATCTAATATCTTGTGATAGTAATAAACAATTCATGTCAG 7327
QY 722 AGTCATTCGAAGCAGATATTCTGATATTGTCATTAACAATTTAGTGAATTATCAT 781
Db 7328 AGTCATTCGAAGCAGATATTCTGATATTGTCATTAACAATTTAGTGAATTATCAT 7387


```
QY 782 CGTCACTTGAATCTGTGTTCAATACGTTCTTAACCTCTTCATATTTAGAAATGAGGCTGA 841
Db 7388 CGTCACTTGAATCTGTGTTCAATACGTTCTTAACCTCTTCATATTTAGAAATGAGGCTGA 7447
QY 842 TGAAGTCCATATTTGAAAAGTTTTCATCACTACTAGTTTGTGATGAGCTTCAAGCCGA 901
Db 7448 TGAAGTCCATATTTGAAAAGTTTTCATCACTACTAGTTTGTGATGAGCTTCAAGCCGA 7507
QY 902 GTTGTCTTTTCTATCTACTCTCATACAACCAATAAATGCTGAAATGAATTTCTAAGCGGA 961
Db 7508 GTTGTCTTTTCTATCTACTCTCATACAACCAATAAATGCTGAAATGAATTTCTAAGCGGA 7567
QY 962 GATGCGCTAGTGATTTTAACTATGCTGCGAGCATTCTTGAGTCCCAATATATAAAGTATT 1021
Db 7568 GATGCGCTAGTGATTTTAACTATGCTGCGAGCATTCTTGAGTCCCAATATATAAAGTATT 7627
QY 1022 GTGTACCTTTTGTCTGAGTGTGTTCTTTAGAGAGATAAAGATCAATGCACTAA 1081
Db 7628 GTGTACCTTTTGTCTGAGTGTGTTCTTTAGAGAGATAAAGATCAATGCACTAA 7687
QY 1082 ACGAAACTGAAACAAGCGATCGAAATAATCCCTTTGGGATCTTGACTCGATAAGTCTAT 1141
Db 7688 ACGAAACTGAAACAAGCGATCGAAATAATCCCTTTGGGATCTTGACTCGATAAGTCTAT 7747
QY 1142 TATTTTCAGAGAAAAATATTCATTTGTTTCTGGGTGGTGATTTGCCAATCATTTCCAT 1201
Db 7748 TATTTTCAGAGAAAAATATTCATTTGTTTCTGGGTGGTGATTTGCCAATCATTTCCAT 7807
QY 1202 TCAGAAATGTTGTTTACACACACCATTCGCCCGGATAAAAAGCATGATGTTGCTGTGG 1261
Db 7808 TCAGAAATGTTGTTTACACACACCATTCGCCCGGATAAAAAGCATGATGTTGCTGTGG 7867
QY 1262 GCATAGAAATTAACCGTCACTCAAAAGGTATAGTTAAATCACTGAATCCGGAGCACTTT 1321
Db 7868 GCATAGAAATTAACCGTCACTCAAAAGGTATAGTTAAATCACTGAATCCGGAGCACTTT 7927
QY 1322 TTCTATTTAAATGAAAAGTGGAAATCTGACATTTCTGGCAACCATTTAACACACAGTCCGA 1381
Db 7928 TTCTATTTAAATGAAAAGTGGAAATCTGACATTTCTGGCAACCATTTAACACACAGTCCGA 7987
QY 1382 ACTGTCCATGAATTTCTGAAAGAGTTAACCCCTCTAAGTAATGAGGTGTTAAGACGCTTT 1441
Db 7988 ACTGTCCATGAATTTCTGAAAGAGTTAACCCCTCTAAGTAATGAGGTGTTAAGACGCTTT 8047
QY 1442 CATTTTCAATGTGGGTAATCGATTTGGCCATACTACTAATCTCTGAATAGCTTTAAGAA 1501
Db 8048 CATTTTCAATGTGGGTAATCGATTTGGCCATACTACTAATCTCTGAATAGCTTTAAGAA 8107
QY 1502 GGTTATGTTTAAACCATCGCTTAATTTGCTGAGATTAAACATAGTCAATGCTTTTCA 1561
Db 8108 GGTTATGTTTAAACCATCGCTTAATTTGCTGAGATTAAACATAGTCAATGCTTTTCA 8167
QY 1562 CTAAAGAAAAAAACATTTTCAAGGAGTGACTGAATTTTCTATTTAATGAATAGTGC 1621
Db 8168 CTAAAGAAAAAAACATTTTCAAGGAGTGACTGAATTTTCTATTAATGAATAGTGC 8227
QY 1622 TTACTTCTTCTTTTGGACCTACAAAACCAATTTTAACATTTCCGATATCGCATTTTTCAC 1681
Db 8228 TTACTTCTTCTTTTGGACCTACAAAACCAATTTTAACATTTCCGATATCGCATTTTTCAC 8287
QY 1682 CATGCTCATCAAGACAGTAAGATAAACATTTGTAACAAAGGAATAGTCATTTCCAACCAT 1741
Db 8288 CATGCTCATCAAGACAGTAAGATAAACATTTGTAACAAAGGAATAGTCATTTCCAACCAT 8347
QY 1742 CTGCTCGTAGGAATGCTTATTTTCTTCTACTGACAGGAATATACCGCCTCTTCAATAA 1801
Db 8348 CTGCTCGTAGGAATGCTTATTTTCTTCTACTGACAGGAATATACCGCCTCTTCAATAA 8407
QY 1802 CACTAAACTCCAACATATAGTAACCTTAATTTTATTAATAACCGCAATTTATTTGGC 1861
Db 8408 CACTAAACTCCAACATATAGTAACCTTAATTTTATTAATAACCGCAATTTATTTGGC 8467
QY 1862 GGCACACACAGATCTCTCTTTAAGTAACTCTCTATTAACAATAGTTTTCATCTAAAAAT 1921
```

```
Db 8468 GGCACACACAGATCTCTCTTTAAGTAACTCTCTATTAACAAGTTTTCATCTAAAAAT 8527
QY 1922 TAGTAGTATTTGAATTAACGGGGCATCGTATTGTAGTTTTCATATTTAGCTTCTGCTT 1981
Db 8528 TAGTAGTATTTGAATTAACGGGGCATCGTATTGTAGTTTTCATATTTAGCTTCTGCTT 8587
QY 1982 CCTTTTGATAACCCACTGTTATTCATGTTGCATGTTGCACTGTTTATACCAACGATATA 2041
Db 8588 CCTTTTGATAACCCACTGTTATTCATGTTGCATGTTGCACTGTTTATACCAACGATATA 8647
QY 2042 GTCTATTAATGCATATATAGTATCCCGAAGCATTTAGCTTCTGAGGCTTCTGAAGAAGCG 2101
Db 8648 GTCTATTAATGCATATATAGTATCCCGAAGCATTTAGCTTCTGAGGCTTCTGAAGAAGCG 8707
QY 2102 TTTCAAGTACTAAATAAGCCGATAGATAGCCACGGAATTCGTAAGCCATTTTCAATAAGTGT 2161
Db 8708 TTTCAAGTACTAAATAAGCCGATAGATAGCCACGGAATTCGTAAGCCATTTTCAATAAGTGT 8767
QY 2162 TAACTTCCGCTCCTCGCTCATTAACAGACATTTCACTACAGTTATGCGGAAAAGTATGCAT 2221
Db 8768 TAACTTCCGCTCCTCGCTCATTAACAGACATTTCACTACAGTTATGCGGAAAAGTATGCAT 8827
QY 2222 GCTGGGTGTGGGGAAGTGTGGAAGAAAAAGAAAGTCAAGTGCCTGCTTTGACATCACTGCT 2281
Db 8828 GCTGGGTGTGGGGAAGTGTGGAAGAAAAAGAAAGTCAAGTGCCTGCTTTGACATCACTGCT 8887
QY 2282 ATCTTCTTACTGTTATGCAAGTCTAGTGGGTGGCACACA 2322
Db 8888 ATCTTCTTACTGTTATGCAAGTCTAGTGGGTGGCACACA 8928
```

```
RESULT 3
ABQ80306
ID ABQ80306 standard; DNA; 1469 BP.
XX
AC ABQ80306;
XX
XX 27-JUN-2003 (first entry)
XX
DE lambda fragment in plasmid DNA.
XX
KW Primer; PCR; amplify; lambda phage; printer material; insertion; SNP;
KM deletion; single nucleotide polymorphism; sequencing; diagnosis; ss.
XX
OS Synthetic.
XX
PN WO2003027991-A1.
XX
XX 03-APR-2003.
PD
XX
XX 24-SEP-2002; 2002WO-JP009766.
PE
XX
XX 25-SEP-2001; 2001JP-00291249.
PR
XX
PA (DNAB-) DNABFORM KK.
PA (RIKE ) RIKEN KK.
XX
PI Hayashizaki Y;
XX
XX WPI; 2003-354676/33.
DR
XX
XX Printed material useful as a delivery and storage system for oligomer
PT and/or polymer, comprises a support having an oligomer and/or polymer
PT applied on it.
XX
XX Example 1; Page 31-32; 91pp; English.
XX
XX The sequences given in ABQ80304-05 are primers which were used to amplify
CC DNA for use in the material of the invention. The primers amplify a 1377
CC bp portion of lambda phage DNA contained in a plasmid. The amplified
CC sequence (ABQ80306) was attached to the printer material of the invention
CC which comprises at least one support having at least one oligomer and/or
```

CC polymer applied on it. The printed material is useful in a method for
CC synthesising DNA where the product of amplification and/or ligation is
CC cDNA and/or full-length cDNA which is recovered and used for
CC determination of nucleotide insertion/deletion, single nucleotide
CC polymorphism (SNP) and sequencing analysis, in a diagnostic method for
CC determination of nucleotide insertion/deletion, or SNP analysis.
CC Optionally, the cDNA and/or full-length cDNA is useful for the peptide,
CC polypeptide or protein expression. The printed material is useful in
CC research applications, or for providing scientists with oligomer and/or
CC polymers from the printed materials easily and immediately. From the
CC printed material, at least an oligomer and/or polymer can be obtained
CC immediately and directly, without need to make a request for it. The
CC oligomers and/or polymers can be delivered and stored easily with reduced
CC labour and time while eliminating the need to use special equipment or
CC facilities. Thus, the printed material is a quick, efficient and
CC inexpensive sample delivery system

SQ Sequence 1469 BP; 411 A; 300 C; 280 G; 478 T; 0 U; 0 Other;

Query Match 28.5%; Score 661.2; DB 7; Length 1469;
Best Local Similarity 98.8%; Pred. No. 2.2e-157;
Matches 666; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1649 CAATTTTAACATTTCCGATATCGATTTTTCACCATGCTCATCAAGAAGAGATAGATAA 1708
DB |||||
DB 32 CCATGATTAACGAATTCGATATCGATTTTTCACCATGCTCATCAAGAAGAGATAGATAA 91
QY 1709 ACAATGTAAACAAAGGAATAGTATCCAAACCATCTGCTGTAAGAAATGCCCTATTTT 1768
DB |||||
DB 92 ACAATGTAAACAAAGGAATAGTATCCAAACCATCTGCTGTAAGAAATGCCCTATTTT 151
QY 1769 CTACTGACGAATATATACCCGCCCTTTCATTAACACTAACTCAACATATAGTAACCT 1828
DB |||||
DB 152 CTACTGACGAATATATACCCGCCCTTTCATTAACACTAACTCAACATATAGTAACCT 211
QY 1829 TAAATTTTAAATAACCGCAATTTATTTGGCGGCAACAGAGATCTCTTTTAACTT 1888
DB |||||
DB 212 TAAATTTTAAATAACCGCAATTTATTTGGCGGCAACAGAGATCTCTTTTAACTT 271
QY 1889 ACTCTCTATTACATACGTTTTCATCTAAATAATTAGTATGAACTTAAACGGGCATC 1948
DB |||||
DB 272 ACTCTCTATTACATACGTTTTCATCTAAATAATTAGTATGAACTTAAACGGGCATC 331
QY 1949 GATATGTAGTTTTCATATTTAGCTTTCTGCTTTCCTTTGGATTAACCCACTGTTATCAT 2008
DB |||||
DB 332 GATATGTAGTTTTCATATTTAGCTTTCTGCTTTCCTTTGGATTAACCCACTGTTATCAT 391
QY 2009 GTTGCAATGGTGCACTGTTTATTAACAAGCATATAGTCTATTAATGATATATAGTATCGCC 2068
DB |||||
DB 392 GTTGCAATGGTGCACTGTTTATTAACAAGCATATAGTCTATTAATGATATATAGTATCGCC 451
QY 2069 GAACGATTAAGCTCTTCAGGCTTGAAGAAGCGTTTCAAGTACTAATAAGCCGATAGATA 2128
DB |||||
DB 452 GAACGATTAAGCTCTTCAGGCTTGAAGAAGCGTTTCAAGTACTAATAAGCCGATAGATA 511
QY 2129 GCCACGGAATCTGTAAGCCATTTTTCATTAAGTGTAACTTCCGCTCTGCTCATTAACAGA 2188
DB |||||
DB 512 GCCACGGAATCTGTAAGCCATTTTTCATTAAGTGTAACTTCCGCTCTGCTCATTAACAGA 571
QY 2189 CATTCATACTACAGTATGCGGGAAGGTATGATGCTGGGTGGGGAAGTCTGGAAGAA 2248
DB |||||
DB 572 CATTCATACTACAGTATGCGGGAAGGTATGATGCTGGGTGGGGAAGTCTGGAAGAA 631
QY 2249 AAGAAGTCAAGCTGCGTCTTGAACATCACTGCTATCTTCTTACTGTTATGACAGTGTGA 2308
DB |||||
DB 632 AAGAAGTCAAGCTGCGTCTTGAACATCACTGCTATCTTCTTACTGTTATGACAGTGTGA 691
QY 2309 GTGGGTGGACACACA 2322
DB |||||
DB 692 GTGGGTGGACACACA 705

RESULT 4

ACC83020
ID ACC83020 standard; DNA; 1469 BP.
XX AC
XX ACC83020;
XX 27-AUG-2003 (first entry)
DT 27-AUG-2003 (first entry)
XX
DE Bacteriophage lambda DNA, SEQ ID NO:3 used in the invention.
XX
KM Oligomer storage; oligomer delivery; ds.
XX
OS Bacteriophage lambda.
XX
PN WO2003040360-A1.
XX
PD 15-MAY-2003.
XX
PF 05-NOV-2002; 2002WO-JP011492.
XX
PR 05-NOV-2001; 2001JP-00339217.
XX
PA (RIKE) RIKEN KK.
PA (DNAP-) DNAPFORM KK.
PA (HAYA/) HAYASHIZAKI Y.
XX
PI Hayashizaki Y;
XX
DR WPI; 2003-441569/41.
XX
PT New support useful for storing and/or delivering an oligomer and/or
PT polymer applied on support, has oligomer and/or polymer applied on it,
PT and a loose-leaf sheet or a card.
XX
PS Example 1; Page 30-31; 70pp; English.

XX The invention relates to a method for storing and/or delivering an
CC oligomer and/or polymer applied on support. The support has oligomer
CC and/or polymer applied on it, and is in the form of loose-leaf sheet or a
CC card. The support is useful for storing and/or delivering an oligomer
CC and/or polymer applied on it. The present sequence is Bacteriophage
CC lambda DNA used in the exemplification of the invention
XX

SQ Sequence 1469 BP; 411 A; 300 C; 280 G; 478 T; 0 U; 0 Other;

Query Match 28.5%; Score 661.2; DB 7; Length 1469;
Best Local Similarity 98.8%; Pred. No. 2.2e-157;
Matches 666; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1649 CAATTTTAACATTTCCGATATCGATTTTTCACCATGCTCATCAAGAAGAGATAGATAA 1708
DB |||||
DB 32 CCATGATTAACGAATTCGATATCGATTTTTCACCATGCTCATCAAGAAGAGATAGATAA 91
QY 1709 ACAATGTAAACAAAGGAATAGTATCCAAACCATCTGCTGTAAGAAATGCCCTATTTT 1768
DB |||||
DB 92 ACAATGTAAACAAAGGAATAGTATCCAAACCATCTGCTGTAAGAAATGCCCTATTTT 151
QY 1769 CTACTGACGAATATATACCCGCCCTTTCATTAACACTAACTCAACATATAGTAACCT 1828
DB |||||
DB 152 CTACTGACGAATATATACCCGCCCTTTCATTAACACTAACTCAACATATAGTAACCT 211
QY 1829 TAAATTTTAAATAACCGCAATTTATTTGGCGGCAACAGAGATCTCTTTTAACTT 1888
DB |||||
DB 212 TAAATTTTAAATAACCGCAATTTATTTGGCGGCAACAGAGATCTCTTTTAACTT 271
QY 1889 ACTCTCTATTACATACGTTTTCATCTAAATAATTAGTATGAACTTAAACGGGCATC 1948
DB |||||
DB 272 ACTCTCTATTACATACGTTTTCATCTAAATAATTAGTATGAACTTAAACGGGCATC 331
QY 1949 GATATGTAGTTTTCATATTTAGCTTTCTGCTTTCCTTTGGATTAACCCACTGTTATCAT 2008
DB |||||
DB 332 GATATGTAGTTTTCATATTTAGCTTTCTGCTTTCCTTTGGATTAACCCACTGTTATCAT 391
QY 2009 GTTGCAATGGTGCACTGTTTATTAACAAGCATATAGTCTATTAATGATATATAGTATCGCC 2068

Db 392 GTTGCATGGTGCACCTGTTATACCAACGATATAGTCTTAATAGCATATATAGTATGCC 451
QY 2069 GAACGATTAGCTCTTCAGGCTTCTGAAGAAGCGTTTCAAGTACTATAAGCCGATAGATA 2128
Db 452 GAACGATTAGCTCTTCAGGCTTCTGAAGAAGCGTTTCAAGTACTATAAGCCGATAGATA 511
QY 2129 GCCACGGAAGCTGCTAGCCATTTTTCATAGTGTTAAGTTCCTCCCTCGCTCATACAGATA 2188
Db 512 GCCACGGAAGCTGCTAGCCATTTTTCATAGTGTTAAGTTCCTCCCTCGCTCATACAGATA 571
QY 2189 CATTCAGTACAGTTATGGCGGAAAGGTATGCATGCTGGGTGGGGAAGTCCGTGAAGATA 2248
Db 572 CATTCAGTACAGTTATGGCGGAAAGGTATGCATGCTGGGTGGGGAAGTCCGTGAAGATA 631
QY 2249 AAGAAGTCAGCTGCGTGTGACATCACTGCTATCTTCTTACTGTTATGACAGTGGTA 2308
Db 632 AAGAAGTCAGCTGCGTGTGACATCACTGCTATCTTCTTACTGTTATGACAGTGGTA 691
QY 2309 GTGGGTGGCACACA 2322
Db 692 GTGGGTGGCACACA 705

RESULT 5
ACD19000 standard; DNA; 10523 BP.
ID ACD19000;
AC ACD19000;
XX
XX
DT 27-OCT-2003 (revised)
DT 21-AUG-2003 (first entry)
XX
DE E. coli 0157 unique DNA sequence OZID_13.
XX
KW OZID; ds; acute haemorrhagic colitis; haemolytic uraemic syndrome;
KW food poisoning.
XX
OS Escherichia coli; strain 0157:H7.
XX
PN US2003023075-A1.
XX
PD 30-JAN-2003.
XX
PF 01-APR-2002; 2002US-00114170.
XX
PR 04-DEC-1998; 98US-0110955P.
PR 03-DEC-1999; 99US-00453702.
XX
PA (BLAT/) BLATTNER F R.
PA (BURL/) BURLAND V D.
PA (PERN/) PERNA N T.
PA (PLUN/) PLUNKETT G.
PA (WELC/) WELCH R.
XX
PI Blattner FR, Burland VD, Perna NT, Plunkett G, Welch R;
XX
DR WPI; 2003-479497/45.
XX
PT New DNA sequences from Escherichia coli strain 0157:H7, useful for
PT detecting E. coli 0157:H7 in a sample, or in designing diagnostic probes
PT which can be used to distinguish strain 0157:H7 from strain K12 using
PT molecular techniques.
XX
PS Claim 16; SEQ ID NO 13; 33bp; English.
XX
CC The invention relates to an isolated DNA molecule comprising an E. coli
CC strain 0157:H7 sequence selected from a clostridial cytotoxin-like gene,
CC a urease gene cluster, a RTX toxin-like gene cluster, a locus of
CC enterocyte effacement and 2 genes from its associated lymphocytic phage
CC 933W (a putative serine/threonine kinase and a tail fibre gene). E. coli
CC 0157:H7 can cause food poisoning, specifically acute haemorrhagic colitis
CC (which can develop into haemolytic uraemic syndrome). Also included are

CC an isolated DNA molecule comprising a nucleotide sequence identical to at
CC least 25 contiguous nucleotides contained in DNA sequences selected from
CC ACD18988-ACD19242 (being 255 E.coli 0157 DNA sequences which are not
CC found in E.coli K12), a recombinant DNA construction comprising the DNA
CC above and a method for detecting E. coli 0157:H7 (ATCC 43895) in a sample
CC (or distinguishing between 0157 and K12) using a probe derived from one
CC of the 255 sequences. The DNA sequences are useful in detecting E. coli
CC 0157:H7 in a sample, for the early diagnosis of humans and livestock
CC infected with 0157:H7, and in designing diagnostic probes which can be
CC used to distinguish strain 0157:H7 from strain K12 using molecular
CC techniques. The present sequence is one of the 255 E. coli 0157:H7 DNA
CC sequence (termed OZID_1-OZID255). Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the USPTO at
CC seqdata.uspto.gov/sequence.html?docid=20030023075 (Updated on 27-OCT-2003
CC to standardise OS field)
SQ Sequence 10523 BP; 2777 A; 2306 C; 2447 G; 2993 T; 0 U; 0 Other;

Query Match 3.2%; Score 75; DB 8; Length 10523;
Best Local Similarity 80.4%; Pred. No. 2.6e-08;
Matches 115; Conservative 0; Mismatches 20; Indels 8; Gaps 2;
QY 2168 CCGCTCCTCGCTCATTAACAGACATTCACTACAGTTATGGCGAAAGGTATGATGCTGG 2227
Db 10258 CCAATCCTTGCTCATTAACAGGACATTCACTTCACTGATGCGAAGATGATGATGGA 10317
QY 2228 TGTGGGAAGTCTGTAAGAAAGAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 2287
Db 10318 -----GAAGTCGTGAAGAAAGAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 10371
QY 2288 TTAAGTGTATGACAGTCTGAT 2310
Db 10372 --ATTGCTATGCAAGTCTGAT 10392

RESULT 6
ADC01049 standard; DNA; 23480 BP.
ID ADC01049;
AC ADC01049;
XX
XX
DT 04-DEC-2003 (first entry)
XX
DE Enterohaemorrhagic E. coli 0157:H7-specific nucleic acid SEQ ID NO: 1093.
XX
KW ds; gene; enterohaemorrhagic; anti-bacterial.
XX
OS Escherichia coli; 0157:H7.
XX
PN JP2002355074-A.
XX
PD 10-DEC-2002.
XX
PF 24-JAN-2002; 2002JP-00015959.
XX
PR 24-JAN-2001; 2001JP-00112010.
XX
PA (UYTS-) UNIV TSUKUBA.
XX
DR WPI; 2003-451640/43.
XX
PT Enterohemorrhagic Escherichia coli 0157:H7-specific nucleic acid molecule
PT and a polypeptide and its use, a polypeptide, a vector and a host cell.
XX
PS Claim 2; SEQ ID NO 1093; 2067bp; Japanese.
XX
CC The invention relates to a novel enterohaemorrhagic Escherichia coli
CC 0157:H7-specific nucleic acid molecule. A polynucleotide of the invention
CC has anti-bacterial activity. The polypeptide can be used in detection
CC and/or treatment of 0157:H7 infection. The nucleotide sequence of the
CC genome of Enterohaemorrhagic E coli 0157:H7 was determined. The present
CC sequence represents an E. coli 0157:H7-specific nucleic acid of the

CC invention.
XX
SQ Sequence 23480 BP; 6393 A; 4893 C; 5323 G; 6871 T; 0 U; 0 Other;
Query Match 3.2%; Score 75; DB 9; Length 23480;
Best Local Similarity 80.4%; Pred. No. 3.5e-08;
Matches 115; Conservative 0; Mismatches 20; Indels 8; Gaps 2;
QY 2168 CCGCTCTCGCTCATATACAGACATTCATCAGTTATGCGGAAAGTATGCATGCTGG 2227
DB 10265 CCAATCTCTGCTCATATACAGGCATTCATTCATGTCAGTAAAGGTATGCATGCTGGA 10324
QY 2228 TGTGGGGAAGTCGTGAAGAAAGAGTCAAGCTGCGTGTGATCATCACTGCTATCTTC 2287
DB 10325 -----GAAGTCGTGAAGAAAGAGTCAAGCTGCGCGCTGTGCTCAAGCTTATCTTC 10378
QY 2288 TTACTGCTTATGCAAGTCGTAGT 2310
DB 10379 --ATTGGCTATGCAAGTCGTAAT 10399

RESULT 7
ABS78855
ID ABS78855 standard; DNA; 9393 BP.
AC ABS78855;
XX
DT 17-DEC-2002 (first entry)
XX
DE E. coli CFT073 genomic sequence #22.
XX
KW Pathogenic; Escherichia coli CFT073 infection; livestock; pyelonephritis;
KW urinary tract infection; open reading frame; ORF; uropathogenic;
KW antibacterial; atropathic; nephrotropic; gene; ds.
XX
OS Escherichia coli.
XX
PN WO200259320-A2.
XX
PD 01-AUG-2002.
XX
PF 19-OCT-2001; 2001WO-US046833.
XX
PR 19-OCT-2000; 2000US-0242412P.
XX
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Blattner FR, Welch RA, Burland VD;
XX
DR WPI; 2002-691532/74.
XX
PT New DNA sequences of the pathogenic Escherichia coli CFT073 strain,
PT useful for preventing or treating E. coli CFT073 infection in humans or
PT livestock.
XX
PS Claim 1; Page 105-110; 765pp; English.
XX
CC The present invention relates to polynucleotide sequences from the genome
CC of the pathogenic Escherichia coli strain CFT073. Almost all the
CC sequences present in E. coli CFT073 are absent in the previously
CC sequenced laboratory strain K-12. The polynucleotide sequences of the
CC invention are useful for preventing, diagnosing or treating E. coli
CC CFT073 infection in humans or livestock. The polynucleotide sequences are
CC useful for preventing urinary tract infections and pyelonephritis.
CC Likewise, the polypeptides encoded by the different open reading frames
CC (ORF1-5) are useful for generating a vaccine against uropathogenic E.
CC coli strains. ABS78834-ABS79085 represent genomic sequences from E. coli
CC strain CFT073
XX
SQ Sequence 9393 BP; 2671 A; 2034 C; 2362 G; 2322 T; 0 U; 4 Other;
Query Match 3.1%; Score 71.4; DB 6; Length 9393;
Best Local Similarity 71.5%; Pred. No. 2.1e-07;

Matches 108; Conservative 0; Mismatches 41; Indels 2; Gaps 1;
QY 2167 TCCGCTCTCGCTCATATACAGACATTCATCAGTTATGCGGAAAGTATGCATGCTGG 2226
DB 8921 TCCGAGTTCGCTCATATATAGACATTCATCAGTTATAGTAAAGGTATGCATGCTGCT 8980
QY 2227 GTGTGGGGAAGTCGTGAAGAAAGAGTCAAGCTGCGTGTGATCATCACTGCTATCTT 2286
DB 8981 GAGTGGGGAAGTGTGAAGAAAGAGACAGCTGCTTGTGTCGTCACGTTATCTT 9040
QY 2287 CTTACTGCTTATGCAAGTCGTAGTGGTGGC 2317
DB 9041 C--ATTGGCTATGCAAGTCGTAAATACAAAGC 9069

RESULT 8
ADA71938/c
ID ADA71938 standard; DNA; 2000 BP.
XX
AC ADA71938;
XX
DT 20-NOV-2003 (first entry)
XX
DE Rice gene, SEQ ID 5263.
XX
KW Plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds.
XX
OS Oryza sativa.
XX
PN WO200300898-A1.
XX
PD 03-JAN-2003.
XX
PF 22-JUN-2001; 2001WO-IB001105.
XX
PR 22-JUN-2001; 2001WO-IB001105.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Chang H, Chen W, Cooper B, Glazebrook U, Golf SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX
DR WPI; 2003-175290/17.
XX
PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX
PS Claim 27; SEQ ID NO 5263; 899pp; English.
XX
CC The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX
SQ Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;
Query Match 2.2%; Score 50.8; DB 7; Length 2000;
Best Local Similarity 12.0%; Pred. No. 0.02;
Matches 43; Conservative 164; Mismatches 151; Indels 0; Gaps 0;
QY 748 ATATTGCTATAAACAATTAGTGAATTATCATCGTCCACTTGAATCTGTGTTCAATTA 807
DB 1014 RTYTRSMSTWYAMKKYTKYMTAYSSSTWKMWAYKRAYAWSRSRKTWCTGGRMATYTC 955

QY 808 CGTCTTAACCTCTTCATATTGAATGAGCGTGAAGTTCATATTGAAAAGTTTCA 867
Db 954 GTKMWAAGRWWRMAMCWCCMWKWKMTSCMWKXWRTWSCWYTMWGMARYAYAMRR 895
QY 868 TCACTACTAGTTTTTGTAGTACCTTCAAGCCAGAGTTGCTTTTCTATCTACTCTATA 927
Db 894 RRWTYKMSWRMRMYMTWKAMTWMTCMCAKWMYMTGMATWMMWRMYTMYTCYAMTCAKC 835
QY 928 CAACCAATAATGCTGAATGTAATTTCTAGCGGAGATCGCCTAGTATTTTAACTATTG 987
Db 834 KYKMAATKMTWTWACAMKATSWRRAMAGMRWKRYKMKRAYWWRWRCWKAGNARMKSR 775
QY 988 CTGGCAGCATTTCTGAGTCCAAATATAAAGTATTGTGTAACCTTTTGTCTGAGTTGT 1047
Db 774 YRWKWKKYATRYWKMAMTWMSWRWKSYRMWSGMRMRMSAMRYCSRMKCAKTKYAS 715
QY 1048 TCTTTAGGAGAGTAAGATCAATGCACTAAACGAACTGAAACAAGCGATCGAA 1105
Db 714 SARWTKRAKRSYRRRMYWKRGWTTYRYWRS CRMTRARMSKRKRWAGASMKSCMW 657

RESULT 9
ADA71938
ID ADA71938 standard; DNA; 2000 BP.
XX

ADA71938;

DT 20-NOV-2003 (first entry)

DE Rice gene, SEQ ID 5263.

XX Plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds.
XX

OS Oryza sativa.

XX WO2003000898-A1.

XX 03-JAN-2003.

XX 22-JUN-2001; 2001WO-IB001105.

XX 22-JUN-2001; 2001WO-IB001105.

PA (SYGN) SYNGENTA PARTICIPATIONS AG.

PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX

DR WPI; 2003-175290/17.

XX Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.

PS Claim 27; SEQ ID NO 5263; 899pp; English.

XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.

XX Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;

Query Match 2.1%; Score 49.8; DB 7; Length 2000;
Best Local Similarity 8.7%; Pred. No. 0.036;

Matches 58; Conservative 309; Mismatches 296; Indels 6; Gaps 1;
QY 1115 TTGGATTCCTTGACTCGATAAGTCTATTATTTCAGAGAAAAATATTCATTCGTTCTG 1174
Db 241 TTSYSWACSSYTWCSKRSRSMWMMKMRKMRWSYGYWSYKMMMTAYKKYSRWCY 300
QY 1175 GGTGCTGATTGCACCAATCATTTCCATTCAAAATTTGTTTACCAACCCATCCGCC 1234
Db 301 MYRGGWRGATRYWGRGYMSRMMYKMYWRYGKMKRGWAGRMMSRMSKACY 360
QY 1235 CGATAAAAGCATGAATGTTCGTGCTGGCATAGATTAAACCGTCACCTCAAAAGTATAG 1294
Db 361 YMRMRMRTRRRRRAKSSRTSRKKRKCWKMRKRYKMRGYSRMRSCRRARMRKCRS 420
QY 1295 TTAATCACTGAATCCGGAGACATTTTCTATTAAATGAAGAAAGTGAAATCTGACAAATT 1354
Db 421 GRAMKMGCRGCMTCRMKSYGMMRWKSWKRMASKXKWSRMYRBRKCKSRTTWMGKTRGG 480
QY 1355 CTGGCAAAACCATTTTACACACGTCGGAACCTGTCATGAATTTCTGAAAGATTACCCCTC 1414
Db 481 MMGTMGRCRYKKRSRSGMKRCKRRRRWGRMY-----RMRMKRYMSARYTMRRCARKKYSY 534
QY 1415 TAACTAATGAGTGTTAAGACCGCTTTCATTTTCAATGTCGGCTAATCGATTGGCCATA 1474
Db 535 SAARKARCWYRGKYWAGWMMKRYKRMWYKMMWYKRYKYSKCSWYCKMSYASCMKSA 594
QY 1475 CTACTAATCCTGGAATAGCTTTAAGAAGTTATGTTTAAACCATCGCTTAATTGCTGA 1534
Db 595 RKAAGMKCRSKMSAWSKMSRSSRCKRCKASKRSSAKRYAMMGMTSGSRMSRWKSYTCY 654
QY 1535 GATTAACATAGTAGTCAATGCTTTTCACTAAGGAAAAAACATTTCAGGAGTTGACTGA 1594
Db 655 WRKWSMKSTCTWMMYMSKYTYAKYGSYWRXYRAMCYMWRWYRYRYSMTYMAWYTS 714
QY 1595 ATTTTATCTATTAAAGAATAAGTGCCTTACTTCTTTTGAACCTACAAAACCAATT 1654
Db 715 STRMAMTGMKYSGRYWTSMYKCKCSWKYRSMWYMSWWAAKTWKMRRYATRMWMMYR 774
QY 1655 TAACATTTCCGATATCGCATTTTTCACCATGCTCAACAGACAGTAAGATAAACATTG 1714
Db 775 YSMKWYTCTWGYMWMYWMRYTKMRMYMWYKCTKYWYWSATYWTGTWAAWMMMAKTKMRM 834
QY 1715 TAACAAAGGAATAGTCAATCCAAACCATCTGCTCGTAGAATGCCTTATTTTCTACTG 1774
Db 835 GMTGAKTRGRARKARYWMMWATWCATKRWMTKGAKAMWMTWAKAMRYWWSWRAMWY 894
QY 1775 CAGGAATAT 1783
Db 895 YYKTRRTRY 903

RESULT 10

ABL56201
ID ABL56201 standard; DNA; 50000 BP.

XX ABL56201;

DT 01-JUL-2002 (first entry)

DE AMEPV genome fragment#3.

XX AMEPV; gene therapy; viral vector; chromosome mapping; gene mapping;
KW genetic deficiency disorder; ds.
XX

OS Amsacta moorei entomopoxvirus.

XX WO200212526-A2.

XX 14-FEB-2002.

XX 10-AUG-2001; 2001WO-US025287.

XX 10-AUG-2000; 2000US-0224479P.

PR 14-SEP-2000; 2000US-00662254.
XX
XX (UYFL) UNIV FLORIDA.
XX
PI Moyer RW, Li Y, Bawden AL;
XX
DR WPI, 2002-227161/28.
XX
PT Novel recombinant entomopox virus vector useful for delivering
PT polynucleotide encoding protein to vertebrate cell, comprises
PT polynucleotide encoding protein operably linked with heterologous
PT promoter sequence.
XX
PS Disclosure; Page 175-201; 326pp; English.
XX
CC The invention relates to a recombinant entomopox virus (EPV) vector,
CC comprising a polynucleotide encoding a protein operably linked with a
CC heterologous promoter sequence. The invention also concerns methods for
CC providing gene therapy for genetic deficiency disorders. Vectors of the
CC invention are useful for delivering a polynucleotide encoding a protein
CC to a vertebrate cell preferably a mammalian cell, such as a human cell.
CC The vector is introduced into the vertebrate cell by infection in a viral
CC particle, or by transfection, transduction, or injection either in vitro
CC or in vivo. The vector is useful for the delivery and expression of
CC biologically useful proteins in gene therapy protocols, and for
CC delivering large DNA segments for engineering of vertebrate cells.
CC Polynucleotides of the invention have applications in techniques such as
CC their use as insertion sites for foreign genes of interest, hybridisation
CC probes, for chromosome and gene mapping, in PCR technologies, and in the
CC production of sense or antisense nucleic acids. Vectors of the invention
CC provide for stable integration and expression of heterologous DNA in host
CC cells, and are adapted for accepting large heterologous polynucleotide
CC inserts which can be delivered in an infected or transformed cell and
CC expressed in a stable fraction. The current sequence represents a
CC fragment of the genome of the genus B entomopoxvirus from amsacta moorei
CC (AMEPV)
XX
SQ Sequence 50000 BP; 19754 A; 4704 C; 4365 G; 21177 T; 0 U; 0 Other;
Query Match 2.0%; Score 45.4; DB 6; Length 50000;
Best Local Similarity 49.5%; Pred. No. 1.5;
Matches 146; Conservative 0; Mismatches 146; Indels 3; Gaps 1;
QY 1235 CGATAAAGCATGAATGTTCTGCTGGGCATAGATTAAACCGTCACCTCAAAAGGTATAG 1294
Db 18484 CGATAAATTCATAATTTTCGATAGACATATAGAACAACTCAACACTTCGATAGTTTGG 18543
QY 1295 TTAATCACTGAATCCGGGACACTTTTCTATTAAATGAAAAGTGAATCTGACAATT 1354
Db 18544 GTAATATATATTTCCATTATATTACAAAACCTCATTTTGAATATATTCAAAATTAAATCAAA 18603
QY 1355 CTGGCAAAACCATTTAACAACGTCGCACTGTCGAATTTCTGAAAAGATTACCCCTC 1414
Db 18604 TTCTAAATATATGTAATAAATTACATTATCTAATCTAAATCTGAATATTTAATATT 18663
QY 1415 TAAGTAATGAGGTGTTAAGACGCTTTCATTTTCAATGTCGGCTAATCGATTGGCCATA 1474
Db 18664 TTATACATTTTGGTATATGATATATATATTATTAATGTAATATAT---GCAATT 18720
QY 1475 CTACTAATCCGTGAATAGCTTTAAGAGGTTATGTTTAAACCATCGCTTAATT 1529
Db 18721 CTCTAAATTTTGTAAATCTTTTAAATTTTCAATATCTATATATCTTTCACATTT 18775
RESULT 11
ABL79740/c
ID ABL79740 standard; cDNA; 453 BP.
XX
AC ABL79740;
XX
DT 17-MAY-2002 (first entry)
XX
DE Human ovarian cancer related cdna clone SEQ ID NO:2718.

XX
KW Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200192581-A2.
XX
PD 06-DEC-2001.
XX
PF 29-MAY-2001; 2001WO-US017756.
XX
PR 26-MAY-2000; 2000US-0207484P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Algate PA, Harlocker SL, Jones R;
XX
DR WPI; 2002-122075/16.
XX
PT Composition for therapy and diagnosis of ovarian cancer comprising
PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
PT polypeptide, antibody specific to polypeptide or T cell expressing
PT polypeptide.
XX
PS Claim 1; SEQ ID NO 2718; 489pp; English.
XX
CC The present invention describes a composition (I) comprising: carriers
CC and immunostimulants; and a polypeptide (II) of a ovarian tumour
CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1)
CC from the 10912 nucleotide sequences as given in ABL77023 to ABL87934',
CC (III) encoding (II) having a sequence (S2), a T cell population of (II),
CC or antigen presenting cells that express (II). (I) has cytostatic
CC activity. An oligonucleotide (IV) that hybridises to (S1) can be used for
CC detecting ovarian cancer in a patient's biological sample preferably
CC serum or ovarian tissue. The method comprises contacting a biological
CC sample from a patient with (IV), detecting the amount of polynucleotide
CC hybridising to (IV) and comparing the amount to a predetermined cutoff
CC value and thereby detecting ovarian cancer in the patient, where the
CC amount of polynucleotide hybridising to (IV) is detected preferably by
CC polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is
CC useful for stimulating and/or expanding T cells specific for an ovarian
CC tumour protein comprising contacting T cells with (III) or (II). (III) is
CC useful in design and preparation of ribozyme molecules for inhibiting
CC expression of the tumour polypeptides and proteins in tumour cells; and
CC to isolate a full length gene from a suitable library e.g., a tumour cDNA
CC library using well known techniques
XX
SQ Sequence 453 BP; 142 A; 63 C; 64 G; 184 T; 0 U; 0 Other;
Query Match 1.9%; Score 44.2; DB 6; Length 453;
Best Local Similarity 46.3%; Pred. No. 0.54;
Matches 145; Conservative 0; Mismatches 168; Indels 0; Gaps 0;
QY 1063 AAAGATCAATGCACTAAACGAAACTGAAACGATCGAAATATCCCTTGGGATT 1122
Db 451 AACACAGCAAGACTCATCTCCAAAAAGACAAAAGAACTAAATATTCTAAGTGCAGAA 392
QY 1123 CTGACTCGATAAGTCTATTATTTTCAAGAAAAAATATTCTATTGTTCTGGGTGGTG 1182
Db 391 CATGATTAGGTAAATATCTCCAAAGCAGAAAGACAGTTTCATATTTCGTTAGTTGAG 332
QY 1183 ATTGACCAATCATTCATTTCAAAATGTTGTTTACACACACCATTCCGCCGATAAAA 1242
Db 331 TCAGTCCCTTCCAAATCAATCTGTTTATTATTAATACAGATGATAGCCAGTAAGT 272
QY 1243 GCATGAATGTTGCTGGGCATAGATTAAACCGTCACCTCAAAAGGTATAGTTAAATCA 1302
Db 271 AATGAGAGAGCAGTCTTTTAAAGCCGATCCATTCTTAAATGAAAAAATATATAATATT 212
QY 1303 CTGAATCCGGGACACTTTTCTATTAAATGAAAAAGTGAAATCTGACAAATCTGGCAA 1362
Db 211 TAGAATAAATTTAATTAATTTCTAAAGTTGAGAAATTTTAAATTTGGGAAA 152

QY 1363 CCATTTAACACAC 1375
Db 151 ATATTTAAACAC 139

RESULT 12
AAK78745/c
ID AAK78745 standard; DNA; 5739 BP.

XX AAK78745;

DT 07-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:33557.

DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytotstatic; gene therapy; vaccine; metastasis; ds.

XX Homo sapiens.

XX WO200157182-A2.

XX PD 09-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001354.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 08-SEP-2000; 2000US-0231413P.

PR 08-SEP-2000; 2000US-0231414P.

PR 08-SEP-2000; 2000US-0232080P.

PR 08-SEP-2000; 2000US-0232081P.

PR 12-SEP-2000; 2000US-0231968P.

PR 14-SEP-2000; 2000US-0232397P.

PR 14-SEP-2000; 2000US-0232398P.

PR 14-SEP-2000; 2000US-0232399P.

PR 14-SEP-2000; 2000US-0232400P.

PR 14-SEP-2000; 2000US-0232401P.

PR 14-SEP-2000; 2000US-0233064P.

PR 14-SEP-2000; 2000US-0233065P.

PR 21-SEP-2000; 2000US-0234223P.

PR 21-SEP-2000; 2000US-0234274P.

PR 25-SEP-2000; 2000US-0234997P.

PR 25-SEP-2000; 2000US-0234998P.

PR 26-SEP-2000; 2000US-0235484P.

PR 27-SEP-2000; 2000US-0235834P.

PR 27-SEP-2000; 2000US-0235836P.

PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236367P.

PR 29-SEP-2000; 2000US-0236368P.

PR 29-SEP-2000; 2000US-0236369P.

PR 02-OCT-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.

PR 02-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-0239935P.

PR 13-OCT-2000; 2000US-0239937P.

PR 20-OCT-2000; 2000US-0240960P.

PR 20-OCT-2000; 2000US-0241221P.

PR 20-OCT-2000; 2000US-0241785P.

PR 20-OCT-2000; 2000US-0241786P.

PR 20-OCT-2000; 2000US-0241787P.

PR 20-OCT-2000; 2000US-0241808P.

PR 20-OCT-2000; 2000US-0241809P.

PR 20-OCT-2000; 2000US-0241826P.

PR 01-NOV-2000; 2000US-0244617P.

PR 08-NOV-2000; 2000US-0246474P.

PR 08-NOV-2000; 2000US-0246475P.

PR 08-NOV-2000; 2000US-0246476P.

PR 08-NOV-2000; 2000US-0246477P.

PR 08-NOV-2000; 2000US-0246478P.

PR 08-NOV-2000; 2000US-0246523P.

PR 08-NOV-2000; 2000US-0246524P.

PR 08-NOV-2000; 2000US-0246525P.

PR 08-NOV-2000; 2000US-0246526P.

PR 08-NOV-2000; 2000US-0246527P.

PR 08-NOV-2000; 2000US-0246528P.

PR 08-NOV-2000; 2000US-0246532P.

PR 08-NOV-2000; 2000US-0246609P.

PR 08-NOV-2000; 2000US-0246610P.

PR 08-NOV-2000; 2000US-0246611P.

PR 08-NOV-2000; 2000US-0246613P.

PR 17-NOV-2000; 2000US-0249207P.

PR 17-NOV-2000; 2000US-0249208P.

PR 17-NOV-2000; 2000US-0249209P.

PR 17-NOV-2000; 2000US-0249210P.

PR 17-NOV-2000; 2000US-0249211P.

PR 17-NOV-2000; 2000US-0249212P.

PR 17-NOV-2000; 2000US-0249213P.

PR 17-NOV-2000; 2000US-0249214P.

PR 17-NOV-2000; 2000US-0249215P.

PR 17-NOV-2000; 2000US-0249216P.

PR 17-NOV-2000; 2000US-0249217P.

PR 17-NOV-2000; 2000US-0249218P.

PR 17-NOV-2000; 2000US-0249244P.

PR 17-NOV-2000; 2000US-0249245P.

PR 17-NOV-2000; 2000US-0249264P.

PR 17-NOV-2000; 2000US-0249265P.

PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX PS Disclosure; SEQ ID NO 33557; 3071bp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
XX SQ Sequence 5739 BP; 1673 A; 1032 C; 1116 G; 1918 T; 0 U; 0 Other;

Query Match 1.9%; Score 44.2; DB 4; Length 5739;
Best Local Similarity 46.3%; Pred. No. 1.4;
Matches 145; Conservative 0; Mismatches 168; Indels 0; Gaps 0;

QY 1063 AAAGATCAATGCCTAAACGAACTGAACAGGATCGAAATATCCCTTGGATT 1122
DB 4336 AACACAGCAAGACCTCTCTCAAAAAAGACAAAGAACTAAATTACTACGAGAA 4277
QY 1123 CTTGACTCGATAAGTCTATTATTTTCAGAGAAAAATATTCATTGTTTGGGTGTG 1182
DB 4276 CATGATTAGGTAATATCTCCAAAGCAGAAAGACAGTTTCATATTTTCGTTAGTTGAG 4217
QY 1183 ATTGCACCAATCATCTCCATTCAAAATGTTGTTTACACACCCCATTCGCCGATAAAA 1242
DB 4216 TCAGTCTTCCTCAAAATCAATCTGTTTTTTTATAGTATACAGATGGTATAGCCAGTAAAGT 4157
QY 1243 GCATGAATGTTGCTGGCATAGATTAACCGTCACCTCAAAAGGTAAGTTAAATCA 1302
DB 4156 AAATGAGAGAGCAGTCTTTTAAGCCGATCCATTCTTAATGAAAAATATAATAATATT 4097
QY 1303 CTGAATCCGGAGCACTTTTCTATTATAAAGAGAGAGTGAATCTGACAATCTGGCAAA 1362
DB 4096 TAGAATAAATTATAATTCTAAAGTTGTAGAAATTTTAAATTTGGATATTTTGGAAA 4037
QY 1363 CCATTTAACACAC 1375

DB 4036 ATATTTAAACAC 4024
RESULT 13
AAK78744/c
ID AAK78744 standard; DNA; 7416 BP.
XX
XX AAK78744;
XX
XX 07-NOV-2001 (first entry)
XX
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:33556.
XX DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX OS Homo sapiens.
XX
XX PN WO200157182-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 17-JAN-2001; 2001WO-US001354.
XX
XX PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.

Db 5713 ATATTAAACCAC 5701
RESULT 14
AAK78746/c
ID AAK78746 standard; DNA; 33513 BP.
XX
AC AAK78746;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:33558.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ds.
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001354.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 14-AUG-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.

PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.

PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and metastasis.
PS Disclosure; SEQ ID NO 33558; 3071bp + Sequence Listing; English.

XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
XX Sequence 33513 BP; 10336 A; 6612 C; 6676 G; 9889 T; 0 U; 0 Other;

Query Match 1.9%; Score 44.2; DB 4; Length 33513;
Best Local Similarity 46.3%; Pred. No. 2.6;
Matches 145; Conservative 0; Mismatches 168; Indels 0; Gaps 0;
QY 1063 AAAGATCAAAATGCACCTAAACGAACTGAACAGCGATCGAAATATCCCTTGGGATT 1122
DB 6798 AACACAGCAAGACCTCATCTCCAAAAAAGACAAAGAACTAAATTACTGACAGAA 6739
QY 1123 CTTGACTCGATTAAGTCTATTAATTTTCAGAGAAAAATATTCATTGTTTCTGGGTG 1182
DB 6738 CATGATTAGTAATATCTCCAAAGCAGAAAGACAGGTTTCATATTTTCGTTAGTTGAG 6679
QY 1183 ATTGACCAATCATTCATTCAAAATGTTGTTTACACACACCAATCCGCCGATAAAA 1242
DB 6678 TCAGTCTTCCAAATCAAAATCTGTTTATTATTAGTATACAGATGTTAGCCAGTAAGT 6619
QY 1243 GCATGAATGTTGCTGGCATAGAAATTAACCGTCACCTCAAAAGTATAGTTAAATCA 1302
DB 6618 AAATGAGAGCAGTCTTTTAAGCCGATCCATTCTTAATGAATAAATATATATTTT 6559
QY 1303 CTGAATCCGGAGCACTTTTCTATTAATGAATAAGTGAATCTGACAATTTCTGCAAA 1362
DB 6558 TAGAATAAATTTATTAATTTCTAAAGTTGAGAAATTTTAATTTGGGAAA 6499
QY 1363 CCATTTAACAAC 1375
DB 6498 ATATTAAACCAAC 6486

RESULT 15
ABQ69245_14
Continuation (15 of 31) of ABQ69245 from base 1400001 (Listeria innocua DNA sequence #684
WP Sequence split into 31 fragments LOCUS ABQ69245 Accession Abq69245
WP Fragment Name Begin End
WP ABQ69245_00 1 110000
WP ABQ69245_01 100001 210000
WP ABQ69245_02 200001 310000
WP ABQ69245_03 300001 410000
WP ABQ69245_04 400001 510000
WP ABQ69245_05 500001 610000
WP ABQ69245_06 600001 710000
WP ABQ69245_07 700001 810000
WP ABQ69245_08 800001 910000
WP ABQ69245_09 900001 1010000
WP ABQ69245_10 1000001 1110000
WP ABQ69245_11 1100001 1210000
WP ABQ69245_12 1200001 1310000
WP ABQ69245_13 1300001 1410000
WP ABQ69245_14 1400001 1510000
WP ABQ69245_15 1500001 1610000
WP ABQ69245_16 1600001 1710000
WP ABQ69245_17 1700001 1810000
WP ABQ69245_18 1800001 1910000
WP ABQ69245_19 1900001 2010000
WP ABQ69245_20 2000001 2110000
WP ABQ69245_21 2100001 2210000
WP ABQ69245_22 2200001 2310000
WP ABQ69245_23 2300001 2410000
WP ABQ69245_24 2400001 2510000
WP ABQ69245_25 2500001 2610000
WP ABQ69245_26 2600001 2710000
WP ABQ69245_27 2700001 2810000
WP ABQ69245_28 2800001 2910000
WP ABQ69245_29 2900001 3010000
WP ABQ69245_30 3000001 3011208

Query Match 1.9%; Score 43.6; DB 6; Length 110000;
Best Local Similarity 56.2%; Pred. No. 5.8;
Matches 82; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
QY 1104 AAATATCCCTTGGATTCTTGACTCGATTAAGTCTATTAATTTTCAGAGAAAAATATTC 1163
DB 86894 AACAAAGCCAAATTTGACCCCAATCTTGAATCTGCTGTTTCTCCAAATAAAGCTAT 86953
QY 1164 ATGTTTCTGGGTTGATGATGACCAATCATTCATTCAAAATGTTGTTTACACACA 1223
DB 86954 ACTTCTTGTGTTTGTGAGTACGCCAAGATTAAATTTTAACATAGTTGATTACCTGA 87013
QY 1224 CCCATTCCGCCGATAAAGCATGAA 1249
DB 87014 GCCGTTAGGGCCGATTAACCAAGTAA 87039

Search completed: October 6, 2004, 05:26:37
Job time : 909 secs

10/674744

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 6, 2004, 05:01:45 ; Search time 9045 Seconds
(without alignments)
11126.868 Million cell updates/sec

Title: US-10-074-744-4
Perfect score: 2322
Sequence: 1 aagcttcttaatttaacct.....gtcgtagtgggtgcacaca 2322

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	2322	100.0	48502	7	LAMCG J02459 Bacterioph
2	1369	59.0	7144	12	AF108217 Cloning v
3	1369	59.0	7232	12	AF179893 Cloning v
4	1260	54.3	133888	2	AC010769 Homo sapi
5	947	40.8	42529	12	CVU39284 Cloning vec
6	947	40.8	42530	12	CVU39285 Cloning vec
7	947	40.8	42531	12	CVU39286 Cloning vec
8	947	40.8	42704	12	CVU37692 Cloning vec
9	486.2	20.9	2242	8	RSSTR1 X62334 R. serpentin
10	329.8	14.2	10124	1	AE006063 Pasteurel
11	201.4	8.7	310967	1	AE016869 Pseudomon
12	142.2	6.1	150050	1	ML0672112 Mesorhizo
13	142.2	6.1	348580	1	BX572601 Rhodospheu
14	104.8	4.5	2365	1	SLHAUR X77574 S. lividans
15	75	3.2	10523	6	AR204117 Sequence
16	75	3.2	12670	1	AE005203 Escherich
17	75	3.2	23480	6	BD184783 Nucleic a
18	75	3.2	295741	1	AP002551 Escherich
19	71.4	3.1	2910	7	PH22XISAA M6185 Bacterioph
20	71.4	3.1	9393	6	AX702436 Sequence
21	71.4	3.1	306358	1	AE016759 Escherich
22	65	2.8	492	6	AR376208 Sequence
23	53.6	2.3	11565	8	AF029891 Cryphonoc
24	52.8	2.3	302156	3	AC116977 Dictyoste
25	52.2	2.2	128683	2	AC135793 Oryza sat
26	51.8	2.2	1141	6	AX083744 Sequence
27	50.8	2.2	2000	6	AX655393 Sequence
28	49.8	2.1	2000	6	AX655393 Sequence
29	49	2.1	183700	2	AC068612 Homo sapi
30	49	2.1	187033	9	AC092965 Homo sapi
31	48.4	2.1	1141	6	AX083744 Sequence
32	48.2	2.1	303414	1	AE015938 Clostridi
33	48	2.1	1302	9	AF090946 Homo sapi
34	47.8	2.1	408	6	AX300942 Sequence
35	47.4	2.0	349751	3	PFMAL4P3 AL035476 Plasmodi
36	47.2	2.0	251762	3	AE014851 Plasmodi
37	47.2	2.0	256172	2	AC005139 Plasmodi
38	47.2	2.0	300050	1	AP004171 Mycoplasma
39	47.2	2.0	310779	2	AC005140 Plasmodi
40	46.8	2.0	721	3	AMU35790 Anopheles m
41	46.8	2.0	55247	9	AL136134 Human DNA
42	46.8	2.0	115859	8	CNS08CE0 BX000510 Oryza sat
43	46.8	2.0	212685	5	BX294380 Zebrafish
44	46.6	2.0	2661	3	AF030966 Schistosom
45	46.4	2.0	140777	8	AP005783 Oryza sat

ALIGNMENTS

RESULT 1
LAMCG LOCUS 48502 bp DNA circular PHG 17-APR-2002
DEFINITION Bacteriophage lambda, complete genome.
ACCESSION J02459 M17233 M24325 V00636 X00906
VERSION J02459.1 GI:215104
KEYWORDS DNA-binding protein; circular; coat protein; complete genome;
origin of replication; repressor; unidentified reading frame.
SOURCE Bacteriophage lambda
ORGANISM Bacteriophage lambda
Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
Lambda-like viruses.
REFERENCE 1 (bases 1 to 12)
AUTHORS Wu, R. and Taylor, E.
TITLE Nucleotide sequence analysis of DNA. II. Complete nucleotide

JOURNAL sequence of the cohesive ends of bacteriophage lambda DNA
MEDLINE J. Mol. Biol. 57 (3), 491-511 (1971)
PUBMED 71209066
4931680
REFERENCE 2 (bases 45493 to 45963)
AUTHORS Imada,M. and Tsugita,A.
TITLE Amino acid sequence of lambda phage endolysin
JOURNAL Nature New Biol. 233, 230-231 (1971)
REFERENCE 3 (sites)
AUTHORS Weigel,P.H., Englund,P.T., Murray,K. and Old,R.W.
TITLE The 3'-terminal nucleotide sequences of bacteriophage lambda DNA
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 70 (4), 1151-1155 (1973)
MEDLINE 73215915
4515613
REFERENCE 4 (bases 38597 to 38672)
AUTHORS Dahlberg,J.E. and Blattner,F.R.
TITLE In vitro transcription products of lambda DNA: Nucleotide sequences and regulatory sites
JOURNAL (in) Fox,C.F. and Robinson,W.S. (Eds.);
VIRUS RESEARCH. PROCEEDINGS OF 1973 ICN-UCLA SYMPOSIUM: 533-544;
Academic Press, New York (1973)
REFERENCE 5 (bases 37945 to 38027)
AUTHORS Maniatis,T., Ptashne,M., Backman,K., Kiehl,D., Flashman,S.,
Jeffrey,A. and Maurer,R.
TITLE Recognition sequences of repressor and polymerase in the operators of bacteriophage lambda
JOURNAL Cell 5 (2), 109-113 (1975)
MEDLINE 75185528
1095210
REFERENCE 6 (bases 35583 to 35600)
AUTHORS Kiehl,D.G., Agarwal,K.L. and Khorana,H.G.
TITLE The nucleotide sequence in the promoter region of the gene N in bacteriophage lambda
JOURNAL J. Biol. Chem. 250 (14), 5574-5582 (1975)
MEDLINE 75189495
167018
REFERENCE 7 (bases 35434 to 35618)
AUTHORS Dahlberg,J.E. and Blattner,F.R.
TITLE Sequence of the promoter-operator proximal region of the major leftward RNA of bacteriophage lambda
JOURNAL Nucleic Acids Res. 2 (9), 1441-1458 (1975)
MEDLINE 76031664
1178525
REFERENCE 8 (bases 37945 to 38018)
AUTHORS Maniatis,T., Jeffrey,A. and Kiehl,D.G.
TITLE Nucleotide sequence of the rightward operator of phage lambda
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 72 (3), 1184-1188 (1975)
MEDLINE 75158212
1055375
REFERENCE 9 (bases 44588 to 44773)
AUTHORS Sklar,J., Yot,P. and Weissman,S.M.
TITLE Determination of genes, restriction sites, and DNA sequences surrounding the 6S RNA template of bacteriophage lambda
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 72 (5), 1817-1821 (1975)
MEDLINE 75217847
1098044
REFERENCE 10 (bases 37905 to 37989)
AUTHORS Walz,,A., Piroletta,V. and Ineichen,K.
TITLE Lambda repressor regulates the switch between PR and P_{rm} promoters
JOURNAL Nature 262 (5570), 665-669 (1976)
MEDLINE 76267718
958438
REFERENCE 11 (bases 37946 to 38039)
AUTHORS Smith,G.R., Eisen,H., Reichardt,L. and Hedgepeth,J.
TITLE Deletions of lambda phage locating a pr_m mutation within the rightward operator
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 73 (3), 712-716 (1976)
MEDLINE 76152323
1062780
REFERENCE 12 (bases 35578 to 35667; 37903 to 38027)
AUTHORS Ptashne,M., Backman,K., Humayun,M.Z., Jeffrey,A., Maurer,R.,
Meyer,B. and Sauer,R.T.
TITLE Autoregulation and function of a repressor in bacteriophage lambda

JOURNAL Science 194 (4261), 156-161 (1976)
MEDLINE 76271154
PUBMED 955943
REFERENCE 13 (bases 35578 to 35667)
AUTHORS Humayun,Z., Jeffrey,A. and Ptashne,M.
TITLE Completed DNA sequences and organization of repressor-binding sites in the operators of phage lambda
JOURNAL J. Mol. Biol. 112 (2), 265-277 (1977)
MEDLINE 77209970
875019
REFERENCE 14 (bases 38610 to 38732)
AUTHORS Scherer,G., Hobom,G. and Kossel,H.
TITLE DNA base sequence of the po promoter region of phage lambda
JOURNAL Nature 265 (5590), 117-121 (1977)
MEDLINE 77100320
834253
REFERENCE 15 (bases 38041 to 38241)
AUTHORS Roberts,T.M., Shimatake,H., Brady,C. and Rosenberg,M.
TITLE Sequence of Cro gene of bacteriophage lambda
JOURNAL Nature 270 (5634), 274-275 (1977)
MEDLINE 78071724
593347
REFERENCE 16 (bases 27616 to 28935)
AUTHORS Davies,R.W., Schreier,P.H. and Buchel,D.E.
TITLE Nucleotide sequence of the attachment site of coliphage lambda
JOURNAL Nature 270 (5639), 757-760 (1977)
MEDLINE 78071823
593399
REFERENCE 17 (bases 37206 to 37263; 37914 to 37970)
AUTHORS Humayun,Z.
TITLE DNA sequence at the end of the ci gene in bacteriophage lambda
JOURNAL Nucleic Acids Res. 4 (7), 2137-2143 (1977)
MEDLINE 78011659
909767
REFERENCE 18 (bases 27617 to 27934)
AUTHORS Landy,A. and Ross,W.
TITLE Viral integration and excision: structure of the lambda att sites
JOURNAL Science 197 (4309), 1147-1160 (1977)
MEDLINE 77258934
331474
REFERENCE 19 (bases 39062 to 39170)
AUTHORS Denniston-Thompson,K., Moore,D.D., Kruger,K.E., Furth,M.E. and Blattner,F.R.
TITLE Physical structure of the replication origin of bacteriophage lambda
JOURNAL Science 198 (4321), 1051-1056 (1977)
MEDLINE 78054731
929187
REFERENCE 20 (bases 44467 to 44807)
AUTHORS Sklar,J.L.
TITLE Structure and function of two regions of DNA controlling the synthesis of prokaryotic RNAs
JOURNAL Thesis (1977)
REFERENCE 21 (sites)
AUTHORS Adhya,S. and Gottesman,M.
TITLE Control of transcription termination
Annu. Rev. Biochem. 47, 967-996 (1978)
MEDLINE 78234064
354508
REFERENCE 22 (bases 13 to 72; 48391 to 48502)
AUTHORS Nichols,B.P. and Donelson,J.E.
TITLE 178-Nucleotide sequence surrounding the cos site of bacteriophage lambda DNA
JOURNAL J. Virol. 26 (2), 429-434 (1978)
MEDLINE 78197067
666898
REFERENCE 23 (bases 37938 to 38016; 35589 to 35666)
AUTHORS Flashman,S.M.
TITLE Mutational analysis of the operators of bacteriophage lambda
Mol. Gen. Genet. 166 (1), 61-73 (1978)
MEDLINE 79114073
368570
REFERENCE 24 (bases 37990 to 38982)

AUTHORS Schwarze,E., Scherer,G., Hobom,G. and Kossel,H.
TITLE Nucleotide sequence of cro, cII and part of the O gene in phage
lambda DNA
JOURNAL Nature 272 (5652), 410-414 (1978)
MEDLINE 78135462
PUBMED 264238
REFERENCE 25 (bases 38212 to 38362)
AUTHORS Rosenberg,M., Court,D., Shimatake,H., Brady,C. and Wulff,D.L.
TITLE The relationship between function and DNA sequence in an
intercistronic regulatory region in phage lambda
JOURNAL Nature 272 (5652), 414-423 (1978)
MEDLINE 78135463
PUBMED 634366
REFERENCE 26 (bases 37224 to 37940)
AUTHORS Sauer,R.T.
TITLE DNA sequence of the bacteriophage gamma ci gene
JOURNAL Nature 276 (5685), 301-302 (1978)
MEDLINE 79053284
PUBMED 714163
REFERENCE 27 (bases 38597 to 39688)
AUTHORS Scherer,G.
TITLE Nucleotide sequence of the O gene and of the origin of replication
in bacteriophage lambda DNA
JOURNAL Nucleic Acids Res. 5 (9), 3141-3156 (1978)
MEDLINE 79033241
PUBMED 704348
REFERENCE 28 (bases 29711 to 29811, 31043 to 31058)
AUTHORS Davies,R.W., Schreier,P.H. and Buchel,D.E.
TITLE Determination of the endpoints of partial deletion mutants of the
attache site of bacteriophage lambda by DNA sequencing
JOURNAL Nucleic Acids Res. 5 (9), 3209-3218 (1978)
MEDLINE 79033246
PUBMED 704352
REFERENCE 29 (bases 21661 to 31129)
AUTHORS Hoess,R.H. and Landy,A.
TITLE Structure of the lambda att sites generated by int-dependent
deletions
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 75 (11), 5437-5441 (1978)
MEDLINE 79074833
PUBMED 364480
REFERENCE 30 (bases 38453 to 38500)

Query Match 100.0%; Score 2322; DB 7; Length 48502;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGCTTTTCTAATTAACTTTGTGAGGTTACCACTACTTAAGTTGTAGGCTCAAGAGG 60
DB 25157 AAGCTTTTCTAATTAACTTTGTGAGGTTACCACTACTTAAGTTGTAGGCTCAAGAGG 25216

QY 61 GTGTGCTCTGTGCTAGTAATAAATACTGACCTGTGAGCTTAATAATTCTATATTGTGTTTC 120
DB 25217 GTGTGCTCTGTGCTAGTAATAAATACTGACCTGTGAGCTTAATAATTCTATATTGTGTTTC 25276

QY 121 TTTCTGCAAAAAAGTGGGAGAGTGAATAAATAATTATTCTAACAATTATCTGCATCAT 180
DB 25277 TTTCTGCAAAAAAGTGGGAGAGTGAATAAATAATTATTCTAACAATTATCTGCATCAT 25336

QY 181 ACCTTCCGAGCATTTTAAAGCATTTCCGCTATAAGTCTCGCTGGAAGAGGTAGTTTTC 240
DB 25337 ACCTTCCGAGCATTTTAAAGCATTTCCGCTATAAGTCTCGCTGGAAGAGGTAGTTTTC 25396

QY 241 CATGTACTTTACCTTCATCTCTGTTCATTTATCATCGCTTTAAACGGCTGCACCTTCT 300
DB 25397 CATGTACTTTACCTTCATCTCTGTTCATTTATCATCGCTTTAAACGGCTGCACCTTCT 25456

QY 301 AATCTATCTGACCATTAATTTTGTAGATGTTTTCATAAGAAAGCTCTGAATCAACG 360
DB 25457 AATCTATCTGACCATTAATTTTGTAGATGTTTTCATAAGAAAGCTCTGAATCAACG 25516

QY 361 GACTGCATATAAGTGTGTATCCAGAAATTTGTCACTTCAAGTAAAAACACCTCACGA 420
DB 25517 GACTGCATATAAGTGTGTATCCAGAAATTTGTCACTTCAAGTAAAAACACCTCACGA 25576

QY 421 GTTAAAAACCTTAAGTCTCAACCGAATGTCTCAATATCCGACGATATATTATTATGCT 480
DB 25577 GTTAAAAACCTTAAGTCTCAACCGAATGTCTCAATATCCGACGATATATTATTATGCT 25636

QY 481 TCTCTGACCGTAGACCTTCCACATGCAGAGATTTTGGAACTCTTGCACTACTACTGGG 540
DB 25637 TCTCTGACCGTAGACCTTCCACATGCAGAGATTTTGGAACTCTTGCACTACTACTGGG 25696

QY 541 GAATGAGTTGCAATTATTGTCTACACCATTTGCGTGCATCGAGTAAAGTCGCTTAATGTTT 600
DB 25697 GAATGAGTTGCAATTATTGTCTACACCATTTGCGTGCATCGAGTAAAGTCGCTTAATGTTT 25756

QY 601 AAAAAAGCAGAGACGAAAGGTGATGCAGATGAACCTCTGTGTTCAATCGAATAAACTAAT 660
DB 25757 AAAAAAGCAGAGACGAAAGGTGATGCAGATGAACCTCTGTGTTCAATCGAATAAACTAAT 25816

QY 661 GACTTTTCCCAACGACATCTAATACTTGTGATAGTAAATAAACAATTGCATGTCCA 720
DB 25817 GACTTTTCCCAACGACATCTAATACTTGTGATAGTAAATAAACAATTGCATGTCCA 25876

QY 721 GAGCTCATTCGAAGCAGATATTCTGATATTGTCTATATAACAATTAGTGAATTTATCA 780
DB 25877 GAGCTCATTCGAAGCAGATATTCTGATATTGTCTATATAACAATTAGTGAATTTATCA 25936

QY 781 TCGTCCACTTGAATCTGTGTTTCATTAACGCTTAACTCTTCATATTAGAATGAGCGT 840
DB 25937 TCGTCCACTTGAATCTGTGTTTCATTAACGCTTAACTCTTCATATTAGAATGAGCGT 25996

QY 841 ATGAGTTCATATTGAAAAGTTTTCATCACTACTACTAGTTTGTGATAGCTTCAAGCCAG 900
DB 25997 ATGAGTTCATATTGAAAAGTTTTCATCACTACTACTAGTTTGTGATAGCTTCAAGCCAG 26056

QY 901 AGTTGTCTTTTCTATCTACTCTCATACACCAATAAATGCTGAATGAATCTAAGCGG 960
DB 26057 AGTTGTCTTTTCTATCTACTCTCATACACCAATAAATGCTGAATGAATCTAAGCGG 26116

QY 961 AGATCGCCTAGTGAATTTAACTAATTGCTGGCAGCAATCTGAGTCCAATATATAAAGTAT 1020
DB 26117 AGATCGCCTAGTGAATTTAACTAATTGCTGGCAGCAATCTGAGTCCAATATATAAAGTAT 26176

QY 1021 TGTGTACCTTTTGTGCTGGGTGAGGTTGTTCTTTAGAGAGAGTAAAGATCAAAATGACACTA 1080
DB 26177 TGTGTACCTTTTGTGCTGGGTGAGGTTGTTCTTTAGAGAGAGTAAAGATCAAAATGACACTA 26236

QY 1081 AAGCAAACTGAACAAGCGATCGAAAATATCCCTTTGGGATCTTGACCTCGATAAGTCTA 1140
DB 26237 AAGCAAACTGAACAAGCGATCGAAAATATCCCTTTGGGATCTTGACCTCGATAAGTCTA 26296

QY 1141 TTAATTTCAAGAAAAAATATTCATTTGTTTCTGSGTTGGTGAATGCAACCAATCATTC 1200
DB 26297 TTAATTTCAAGAAAAAATATTCATTTGTTTCTGSGTTGGTGAATGCAACCAATCATTC 26356

QY 1201 TTCAAAATTTGTTTACCAACCCATTCGCGCCGATATAAAGCATGAATGTCGTGCTG 1260
DB 26357 TTCAAAATTTGTTTACCAACCCATTCGCGCCGATATAAAGCATGAATGTCGTGCTG 26416

QY 1261 GGCATAGAAATTAACCGTCACTCAAAAAGGTATAGTTAAATCACTGAATCCGGAGACACTT 1320
DB 26417 GGCATAGAAATTAACCGTCACTCAAAAAGGTATAGTTAAATCACTGAATCCGGAGACACTT 26476

QY 1321 TTTCTATTAATGAAAAGTGAATACTGACAATTTCTGGCAACCATTTAACAACAGTGGC 1380
DB 26477 TTTCTATTAATGAAAAGTGAATACTGACAATTTCTGGCAACCATTTAACAACAGTGGC 26536

QY 1381 AACTGTCCATGAATTTCTGAAGAAGTTAACCCCTCTAAGTAAAGAGTGTTAAGAGCGCTT 1440
DB 26537 AACTGTCCATGAATTTCTGAAGAAGTTAACCCCTCTAAGTAAAGAGTGTTAAGAGCGCTT 26596

QY 1441 TCAATTTCAATGTGGGCTAATTCGATTTGGCCATACTAATAATCCTGAATAGCTTTAAGA 1500
DB 26597 TCAATTTCAATGTGGGCTAATTCGATTTGGCCATACTAATAATCCTGAATAGCTTTAAGA 26656

QY 1501 AGTTATGTTTAAACCATCGCTTAATTGCTGAGATTAACTAGTCAATGCTTTCA 1560
| | | | |
Db 26657 AGGTTATGTTTAAACCATCGCTTAATTGCTGAGATTAACTAGTCAATGCTTTCA 26716
QY 1561 CCTAAGAAAAAACATTTCAGGAGTGAAGTGAATTTTAACTATTAATGAATAAGTG 1620
| | | | |
Db 26717 CCTAAGAAAAAACATTTCAGGAGTGAAGTGAATTTTAACTATTAATGAATAAGTG 26776
QY 1621 CTTAAGTCTTTCTTTTGAAGCTTAAACCAATTTTAACTATTCGATATCGCATTTTCA 1680
| | | | |
Db 26777 CTTAAGTCTTTCTTTTGAAGCTTAAACCAATTTTAACTATTCGATATCGCATTTTCA 26836
QY 1681 CCAAGCTCATCAAGACAGTAAGATTAACCAATTTGTAACAAAGGATAGTCATTTCAACCA 1740
| | | | |
Db 26837 CCAAGCTCATCAAGACAGTAAGATTAACCAATTTGTAACAAAGGATAGTCATTTCAACCA 26896
QY 1741 TCTGCTCGTAGAATGCGCTTATTTTCTACTGACGAATATACCCGCTCTTCAATA 1800
| | | | |
Db 26897 TCTGCTCGTAGAATGCGCTTATTTTCTACTGACGAATATACCCGCTCTTCAATA 26956
QY 1801 ACACCTAACTCCACATATAGTAACCCCTTAATTTTAAATAAACCGCAATTTATTTGG 1860
| | | | |
Db 26957 ACACCTAACTCCACATATAGTAACCCCTTAATTTTAAATAAACCGCAATTTATTTGG 27016
QY 1861 CGGCAACACAGGATCTCTCTTTTAAAGTACTCTCTATTACATACGTTTCCATCTAATAA 1920
| | | | |
Db 27017 CGGCAACACAGGATCTCTCTTTTAAAGTACTCTCTATTACATACGTTTCCATCTAATAA 27076
QY 1921 TTAGTAGTATTTGAACCTTAAACGGGCGATCGTATTGTAGTTTTCATATTAGCTTTCTGCT 1980
| | | | |
Db 27077 TTAGTAGTATTTGAACCTTAAACGGGCGATCGTATTGTAGTTTTCATATTAGCTTTCTGCT 27136
QY 1981 TCCCTTTGGATAACCCACTGTTTATTCATGTTGCATGTCGACTGTTTATACCAACGATAT 2040
| | | | |
Db 27137 TCCCTTTGGATAACCCACTGTTTATTCATGTTGCATGTCGACTGTTTATACCAACGATAT 27196
QY 2041 AGCTAATTAATGCATATATAGTAATGCGCCGAACGATTAGCTCTTCAGGCTTCTGAAGAAGC 2100
| | | | |
Db 27197 AGCTAATTAATGCATATATAGTAATGCGCCGAACGATTAGCTCTTCAGGCTTCTGAAGAAGC 27256
QY 2101 GTTTCAGTACTAATAAGCCGATAGATAGCCACGACTTCGTAGCCATTTTTCATAAGTG 2160
| | | | |
Db 27257 GTTTCAGTACTAATAAGCCGATAGATAGCCACGACTTCGTAGCCATTTTTCATAAGTG 27316
QY 2161 TTAACCTCCGCTCCTCGCTCATATAACAGACATTCACATAGTAATGCGGAAAGGTATGCA 2220
| | | | |
Db 27317 TTAACCTCCGCTCCTCGCTCATATAACAGACATTCACATAGTAATGCGGAAAGGTATGCA 27376
QY 2221 TGTGGGTGTGGGGAAGTCGTGAAGAAAGAAAGAGTCAAGTCGTCGTTTGACATCACTGC 2280
| | | | |
Db 27377 TGTGGGTGTGGGGAAGTCGTGAAGAAAGAAAGAGTCAAGTCGTCGTTTGACATCACTGC 27436
QY 2281 TATCTTCTTACTGTTATGACAGGTCGTAGTGGGTGGCACACA 2322
| | | | |
Db 27437 TATCTTCTTACTGTTATGACAGGTCGTAGTGGGTGGCACACA 27478

RESULT 2
AF108217 7144 bp DNA circular SYN 16-APR-1999
LOCUS AF108217
DEFINITION Cloning vector pLM10, complete sequence.
ACCESSION AF108217
VERSION AF108217.1 GI:4585224
KEYWORDS
SOURCE Cloning vector pLM10
ORGANISM Cloning vector pLM10
artificial sequences; vectors.
REFERENCE 1 (bases 1 to 7144)
AUTHORS Marino-Ramirez,L., Blum,B.J. and Hu,J.C.
TITLE Green fluorescent protein (GFP) as a reporter to map
protein-protein interactions in Escherichia coli
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 7144)

AUTHORS Marino-Ramirez,L., Blum,B.J. and Hu,J.C.
TITLE Direct Submission
JOURNAL Submitted (19-NOV-1998) Biochemistry & Biophysics, Texas A&M
University, College Station, TX 77843-2128, USA
FEATURES
source location/Qualifiers
1. 7144
/organism="Cloning vector pLM10"
/mol_type="other DNA"
/specific_host="Escherichia coli Ag1688"
/db_xref="taxon:89245"
complement(4028..4888)
/gene="bla"
complement(4028..4888)
/gene="bla"
complement(4028..4888)
/gene="bla"
/codon_start=1
/transl_table=11
/product="beta-lactamase"
/protein_id="AAD25334.1"
/db_xref="GI:4585225"
/translation="MSIQHFRVALIPFPAFCIPVFAHPETLVKVKDAEDQLGARVGY
IEDLNSGKILDSFRPEERFPMSTFKVLLCGAVLSRVDAQEQGLRIHYSQNDLVE
YSPVTEKHLTDGMIVRELCSAAITMSDNTAANLLTTIGPKELTAFLHNMGDVTRL
DRWEPELNEAIPNDERDITMPVAMATTLRLKLITGELLTLASRQQLIDWEADKVAEPL
LRSAIPAGWFIADKSGAGERGSRGIIAALGPDKSRIVITIYTTGSQATMDERNRQIA
EIGASLIKHW"
complement(5500..6159)
/gene="cat"
complement(5500..6159)
/gene="cat"
/codon_start=1
/transl_table=11
/product="chloramphenicol acetyltransferase"
/protein_id="AAD25335.1"
/db_xref="GI:4585226"
/translation="MEKITGYTTVDISQWHRKHEFAFQSVAACTYNQTVOLDITAF
LKTVKKNKHKFYPAFLHILARLMAHPEFRAMKDELVIWDSVHPCYTFVHEQTETP
SSLMSBYHDDFRQFLHIVSQDVACYGENLAYFPKGFLENMEFVSANPWVSFTSPDLNV
ANMDFPAPVFETMGKYYTGQDKVLMPLAIQVHVAVDGFHVGRMLNELQQYCDDEWQGG
A"

promoter complement(6212..6269)
/note="lacUV5 promoter without operator"
/evidence=experimental
6276..7132
/gene="gfp"
/evidence=experimental
/allele="mut2"
6276..6330
/gene="gfp"
/note="pl promoter"
/evidence=experimental
6281..6295
/gene="gfp"
/bound_molety="O-L1 repressor"
6305..6321
/gene="gfp"
/bound_molety="O-L1 repressor"
6376..7092
/gene="gfp"
/note="optimized for FACS"
/codon_start=1
/transl_table=11
/product="green fluorescent protein"
/protein_id="AAD25336.1"
/db_xref="GI:4585227"
/translation="MSKGEHLFTGVPIPLVELDGVNNGHKFSVSGEGBDATYKLTLL
KFICTTGKLPVPMPTLVTFAYGLQCFARYPDMMKQHDFEKSAMPEGYOERTIFFKD
DGNVKTAEVKEFGDTLVNRIELKGIDFKEDGNILGHKLEYNNSHVYIMADKQNG
IKVNFKIRHNIEDGSVQLADHYQNTPIGDGPVLLPDNNHYLSTQSALSKDPNEKRDM
VLEFVTAAGITGHGMDLYK"
7109..7132
/gene="gfp"

ORIGIN

attenuator

Query Match 59.0%; Score 1369; DB 12; Length 7144;
Best Local Similarity 100.0%; Pred. No. 2.8e-276;
Matches 1369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 954 TAAAGGAGATCGCCTAGTATTTAACTATTGCTGGCAGCATCTTGAGTCCCATATA 1013
DB 1 TAAGCGGAGATCGCCTAGTATTTAACTATTGCTGGCAGCATCTTGAGTCCCATATA 60

QY 1014 AAAGTATTTGTACTTTTGTGGGTGAGGTGTTCTTTAGAGAGATAAAGGATCAAA 1073
DB 61 AAAGTATTTGTACTTTTGTGGGTGAGGTGTTCTTTAGAGAGATAAAGGATCAAA 120

QY 1074 TGCATTAACGAAACTGAAACAGCGATCGAAATATCCCTTTGGGATCTTGACTCGAT 1133
DB 121 TGCATTAACGAAACTGAAACAGCGATCGAAATATCCCTTTGGGATCTTGACTCGAT 180

QY 1134 AAGTCTATTATTTTCAGAGAAAAATATTCATTGTTTCTGGGTGGTGTGATGACCAAT 1193
DB 181 AAGTCTATTATTTTCAGAGAAAAATATTCATTGTTTCTGGGTGGTGTGATGACCAAT 240

QY 1194 CATTCATTCAAATTTGTTTTCACACACCCATTCGGCCGATAAAGCATGAATGTT 1253
DB 241 CATTCATTCAAATTTGTTTTCACACACCCATTCGGCCGATAAAGCATGAATGTT 300

QY 1254 CGTGTGGGCATAGATTAACCCGTCACCTCAAAAGGTATAGTTAAATCACTGAATCCGG 1313
DB 301 CGTGTGGGCATAGATTAACCCGTCACCTCAAAAGGTATAGTTAAATCACTGAATCCGG 360

QY 1314 AGCACTTTTCTATTAAATGAAAAGTGAATCTGACAAATCTGGCAAAACCATTTAACAC 1373
DB 361 AGCACTTTTCTATTAAATGAAAAGTGAATCTGACAAATCTGGCAAAACCATTTAACAC 420

QY 1374 ACGTGGCAACTGTCATGAATTTCTGAAAAGATTACCCCTCTAAGTAATGAGGTGTTAG 1433
DB 421 ACGTGGCAACTGTCATGAATTTCTGAAAAGATTACCCCTCTAAGTAATGAGGTGTTAG 480

QY 1434 GACGCTTTCATTTTCAATGTCGGCTAATCGATTGGCCACTACTAAATCCCTGAATAGC 1493
DB 481 GACGCTTTCATTTTCAATGTCGGCTAATCGATTGGCCACTACTAAATCCCTGAATAGC 540

QY 1494 TTTAAGAAGGTTATGTTTAAACCATCGCTTAATTTGCTGAGATTAACATAGTCAAT 1553
DB 541 TTTAAGAAGGTTATGTTTAAACCATCGCTTAATTTGCTGAGATTAACATAGTCAAT 600

QY 1554 GCTTTCACCTAAGGAAAAACAATTCAGGGAGTTGACTGAATTTTATCTATTATGA 1613
DB 601 GCTTTCACCTAAGGAAAAACAATTCAGGGAGTTGACTGAATTTTATCTATTATGA 660

QY 1614 ATAAGTCTTACTCTCTCTTTTGAACCTACAAAACCAATTTTAACTCCGATATCGCA 1673
DB 661 ATAAGTCTTACTCTCTCTTTTGAACCTACAAAACCAATTTTAACTCCGATATCGCA 720

QY 1674 TTTTTCACCATGCTCATCAAGACAGTAAGATAAAACATTTGTAACAAAGGATAGTCAAT 1733
DB 721 TTTTTCACCATGCTCATCAAGACAGTAAGATAAAACATTTGTAACAAAGGATAGTCAAT 780

QY 1734 CCAACCATCTGCTCGTAGAATGCTTATTTTCTACTGAGGAATATACCCGCTCT 1793
DB 781 CCAACCATCTGCTCGTAGAATGCTTATTTTCTACTGAGGAATATACCCGCTCT 840

QY 1794 TTCAATAACAATACTCCAACATATAGTAACCCCTTAATTTTAAATAACCCGAAT 1853
DB 841 TTCAATAACAATACTCCAACATATAGTAACCCCTTAATTTTAAATAACCCGAAT 900

QY 1854 TATTTGGCGGCAACACAGAGATCTCTTTTAAAGTTACTCTTAATTAACATACGTTTCCAT 1913
DB 901 TATTTGGCGGCAACACAGAGATCTCTTTTAAAGTTACTCTTAATTAACATACGTTTCCAT 960

QY 1914 CTAAAAATTAGTAGTATTGAATTAACGGGGCATCGTATGTAGTTTCCATATTTAGCT 1973
DB 961 CTAAAAATTAGTAGTATTGAATTAACGGGGCATCGTATGTAGTTTCCATATTTAGCT 1020

QY 1974 TTCGTCTCTTTTGGATAACCACTGTATTATCATGTGATGGTGCACTGTTATACCA 2033

DB 1021 TTCGTCTCTTTTGGATAACCACTGTATTATTCATGTGGATGGTGCACTGTTATACCA 1080

QY 2034 ACGATATAGTCTATTAAATGATATATAGTATCGCCGAACGATTAAGCTCTCAGGCTCTG 2093
DB 1081 ACGATATAGTCTATTAAATGATATATAGTATCGCCGAACGATTAAGCTCTCAGGCTCTG 1140

QY 2094 AAGAACGTTTCAAGTACTAATAAGCCGATAGATAGCCAGCACTCGTAGCCATTTTC 2153
DB 1141 AAGAACGTTTCAAGTACTAATAAGCCGATAGATAGCCAGCACTCGTAGCCATTTTC 1200

QY 2154 ATAAGTGTAACTTCCGCTCCTCGCTCATACAGACATTCACACTACAGTTATGGCGGAAG 2213
DB 1201 ATAAGTGTAACTTCCGCTCCTCGCTCATACAGACATTCACACTACAGTTATGGCGGAAG 1260

QY 2214 GTATGCATGCTGGGTGTGGGGAAGTCTGAAAGAAAAGAGTCAAGCTGCGTGTGACA 2273
DB 1261 GTATGCATGCTGGGTGTGGGGAAGTCTGAAAGAAAAGAGTCAAGCTGCGTGTGACA 1320

QY 2274 TCACTGCTATCTTCTTACTGTTATGACAGTCTGATGAGGTGGACACA 2322
DB 1321 TCACTGCTATCTTCTTACTGTTATGACAGTCTGATGAGGTGGACACA 1369

RESULT 3
AF179893 7232 bp DNA circular SYN 05-OCT-1999
LOCUS AF179893
DEFINITION Cloning vector pLM53, complete sequence.
ACCESSION AF179893
VERSION AF179893.1 GI:6010635
KEYWORDS
SOURCE Cloning vector pLM53
ORGANISM Cloning vector pLM53
REFERENCE 1 (bases 1 to 7232)
AUTHORS Marino-Ramirez, L. and Hu, J.C.
TITLE Novel Escherichia coli strains to detect protein-protein interactions
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 7232)
AUTHORS Marino-Ramirez, L. and Hu, J.C.
TITLE Direct Submission
JOURNAL Submitted (23-AUG-1999) Biochemistry and Biophysics, Texas A & M University, College Station, TX 77843-2128, USA
FEATURES
source
1. 7232
/organism="Cloning vector pLM53"
/mol_type="other DNA"
/db_xref="taxon:103609"
/lab_host="Escherichia coli AG1688"
complement(4028..4888)
/gene="bla"
complement(4028..4888)
/gene="bla"
/codon_start=1
/evidence=experimental
/product="beta-lactamase"
/protein_id="AA01199.1"
/db_xref="GI:6010636"
/translation="MSIQHFRVALIPFPAFLPVFAHPETLVKVKDAEDQLGARVGY
IELDINSGLIESFRPERFPMSTFKVLGAVLSRVDAQEOLGRIRIHSQNDLVE
YSPVTEKHLTDGWTYRELCSAII TMSDNTAAILLITIGPEKELTAFILHMGDHYRL
DRMEPELNEALPNDERDITMPVAMATTLRKLLTGLSLTASRQQLIDWMEADKVA
PLRSALPAGWFIADKSGAGERGSRGIIALGPDKSRIVIVITTTGSQATMDERNRQIA
EIGASLIKHW"
complement(5500..6285)
/gene="cat"
complement(5500..6159)
/gene="cat"
/codon_start=1
/evidence=experimental
/transl_table=11
CDS
gene
CDS

/product="chloramphenicol acetyltransferase"
/protein_id="AA01200.1"
/db_xref="GI:6010637"
/translation="MEKITGYTTVDISQWHRKEHFAQSVACQTYNQVQLDITAF
LKTVKKNKHKEYPAPFHLILMNAHPEFRMAMDELIVMSVHPCYTFHEQTEF
SSLMSXYHDFRQFLHIYSQDVACYENLAYFPKGFLENMFVSANPWVSFTSFDLNV
ANMDFNFAFVFTMGKXYTQGDKVLMLAIQVHNAVCDGFHVGRMLNELQYCDIEWQGG
A"
complement(6231. .6285)
/gene="cat"
/note="pl promoter"
/evidence=experimental
complement(6240. .6256)
/gene="cat"
/bound_moiety="O-L1 repressor"
complement(6264. .6280)
/gene="cat"
/bound_moiety="O-L2 repressor"
6586. .7119
/gene="gat"
6586. .7119
/gene="gat"
/codon_start=1
/evidence=experimental
/transl_table=11
/product="gentamicin acetyltransferase"
/protein_id="AA01201.1"
/db_xref="GI:6010638"
/translation="MLRSSNDVTQOGRPKIKLGGSSMGIIIRTCRLGPDQVKSMAAL
DLFGRFEGDVATYSQHDPSDYLGNLRSKTFIALAEPQEAUVGALAAYVLPREFQF
RSEIYLYDLAVSGEHRQGIATALLNLKHEANALGAYIVYQADYGDPPAVALYTKL
GIREVMHFIDIPSTAT"

ORIGIN

Query Match	59.0%;	Score 1369;	DB 12;	Length 7232;
Best Local Similarity	100.0%;	Pred. No. 2.8e-276;		
Matches 1369;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 954	TAAGCGAGATCGCCTAGTATTTAAACTATTGCTGCAGACATTTCTGAGTCCAATATA			1013
Db 1	TAAGCGAGATCGCCTAGTATTTAAACTATTGCTGCAGACATTTCTGAGTCCAATATA			60
QY 1014	AAAGTATGTGTACCTTTGCTGGGTCAAGTGTGCTTTAGAGAGAGTAAAGGATCAAA			1073
Db 61	AAAGTATGTGTACCTTTGCTGGGTCAAGTGTGCTTTAGAGAGAGTAAAGGATCAAA			120
QY 1074	TGCACATAACGAAACTGAACCAAGCATCGAAATATCCCTTTGGGATTTCTGACTCGAT			1133
Db 121	TGCACATAACGAAACTGAACCAAGCATCGAAATATCCCTTTGGGATTTCTGACTCGAT			180
QY 1134	AAAGTATTTATTTTCAGAGAAAAATATTCATTGTTTCTGGGTGGTGATTCACCAAT			1193
Db 181	AAAGTATTTATTTTCAGAGAAAAATATTCATTGTTTCTGGGTGGTGATTCACCAAT			240
QY 1194	CATTCCATTCAAATTTGTTGTTTACACACACCCCATTCGCGCCGATAAAAGCATGAATGT			1253
Db 241	CATTCCATTCAAATTTGTTGTTTACACACACCCCATTCGCGCCGATAAAAGCATGAATGT			300
QY 1254	CGTGTGGGCATAGAAATTAACCGGTACCTCAAAAAGGTATAGTTAAATCACTGAATCCGGG			1313
Db 301	CGTGTGGGCATAGAAATTAACCGGTACCTCAAAAAGGTATAGTTAAATCACTGAATCCGGG			360
QY 1314	AGCACTTTTCTATTAAATGAAGAAGTGAATCTGACAATTCGCAAAACCATTTAACAC			1373
Db 361	AGCACTTTTCTATTAAATGAAGAAGTGAATCTGACAATTCGCAAAACCATTTAACAC			420
QY 1374	ACGTGCAACTGTCCATGAATTTCTGAAGAAGTTACCCCTCTAAGTAATGAGGTGTTAG			1433
Db 421	ACGTGCAACTGTCCATGAATTTCTGAAGAAGTTACCCCTCTAAGTAATGAGGTGTTAG			480
QY 1434	GACGCTTTCATTTTCATATGTCGGCTAATCGATTGGCCATACTATAAATCCTGAATAGC			1493
Db 481	GACGCTTTCATTTTCATATGTCGGCTAATCGATTGGCCATACTATAAATCCTGAATAGC			540

QY 1494	TTTAAGAAGGTATATGTTTAAACCATCGCTTAATTGCTGAGATTACATAGTAGCAAT	1553
Db 541	TTTAAGAAGGTATATGTTTAAACCATCGCTTAATTGCTGAGATTACATAGTAGCAAT	600
QY 1554	GCITTCACCTAAGGAAAAAACAATTTCAAGGAGTGAAGTATTTTATCTATTAATGA	1613
Db 601	GCITTCACCTAAGGAAAAAACAATTTCAAGGAGTGAAGTATTTTATCTATTAATGA	660
QY 1614	ATAAGTCTTACTTCTCTTTTGAACCTACAAAACCAATTTTAACATTTCCGATATCGCA	1673
Db 661	ATAAGTCTTACTTCTCTTTTGAACCTACAAAACCAATTTTAACATTTCCGATATCGCA	720
QY 1674	TTTTTCACCATGCTCATCAAGACAGTAAGATAAACAATTGTAAACAAGGATAGTCATT	1733
Db 721	TTTTTCACCATGCTCATCAAGACAGTAAGATAAACAATTGTAAACAAGGATAGTCATT	780
QY 1734	CCAACCATCTGCTCGTAGGAATGCTTATTTTCTACTGACAGAAATATACCCGCTCT	1793
Db 781	CCAACCATCTGCTCGTAGGAATGCTTATTTTCTACTGACAGAAATATACCCGCTCT	840
QY 1794	TTCAATAACACTAACTCCAACATATAGTAACCCCTTAATTTTAAATAACCGCAATT	1853
Db 841	TTCAATAACACTAACTCCAACATATAGTAACCCCTTAATTTTAAATAACCGCAATT	900
QY 1854	TATTTGGCGGCAACACAGGATCTCTTTTAAGTACTCTCTATTAACATACGTTTCCAT	1913
Db 901	TATTTGGCGGCAACACAGGATCTCTTTTAAGTACTCTCTATTAACATACGTTTCCAT	960
QY 1914	CTAATAATTAGTAGTATTTGAACCTTAACGGGGCATCGATTTGTAGTTTTCATATTTAGCT	1973
Db 961	CTAATAATTAGTAGTATTTGAACCTTAACGGGGCATCGATTTGTAGTTTTCATATTTAGCT	1020
QY 1974	TTCTGCTTCTTTTGGATAACCACTGTATTATCATGTGTCAGTGTATATACCA	2033
Db 1021	TTCTGCTTCTTTTGGATAACCACTGTATTATCATGTGTCAGTGTATATACCA	1080
QY 2034	ACGATATAGTCTATTAAATGCAATATATAGTATCGCCGAACGATTAGCTCTCAGGCTTCTG	2093
Db 1081	ACGATATAGTCTATTAAATGCAATATATAGTATCGCCGAACGATTAGCTCTCAGGCTTCTG	1140
QY 2094	AAAGAGCTTTCAAGTACTAATAAGCCGATAGATAGCCACGCACTTCGTAGCCATTTTTC	2153
Db 1141	AAAGAGCTTTCAAGTACTAATAAGCCGATAGATAGCCACGCACTTCGTAGCCATTTTTC	1200
QY 2154	ATAAGTGTAACTTCCGCTCCTCGCTCATTAACAGACATTCACAGTTATGGCGGAAG	2213
Db 1201	ATAAGTGTAACTTCCGCTCCTCGCTCATTAACAGACATTCACAGTTATGGCGGAAG	1260
QY 2214	GTAATGATGCTGGGTGTGGGGAAGTCGTGAAGAAGAAAGAGTACGCTGCGTGTGACA	2273
Db 1261	GTAATGATGCTGGGTGTGGGGAAGTCGTGAAGAAGAAAGAGTACGCTGCGTGTGACA	1320
QY 2274	TCACTGCTATCTTCTTACTGTTATGCAAGTGTAGTGGGTGGCACACA	2322
Db 1321	TCACTGCTATCTTCTTACTGTTATGCAAGTGTAGTGGGTGGCACACA	1369

RESULT 4
AC010769/c 133888 bp DNA linear HTG 07-MAY-2001
LOCUS Homo sapiens chromosome 15 clone RP11-28B17 map 15, WORKING DRAFT
DEFINITION SEQUENCE, 11 unordered pieces.
ACCESSION AC010769
VERSION AC010769.3 GI:7341761
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 133888)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 15, clone RP11-28B17

JOURNAL
REFERENCE
AUTHORS

Unpublished
2 (bases 1 to 133888)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczky,J., Lieu,C., Locke,K., MacDonald,P., Margulis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

TITLE
JOURNAL

Submitted (22-SEP-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 30, 2000 this sequence version replaced gi:6721434.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: l2251

Center clone name: 28_B_17

Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 127855 bases at least Q40

Consensus quality: 131192 bases at least Q30

Consensus quality: 132364 bases at least Q20

Insert size: 121000; agarose-fp

Insert size: 132888; sum-of-contigs

Quality coverage: 6.6 in Q20 bases; agarose-fp

Quality coverage: 6.0 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1054: contig of 1054 bp in length
* 1055 1154: gap of 100 bp
* 1155 2415: contig of 1261 bp in length
* 2416 2515: gap of 100 bp
* 2516 3643: contig of 1128 bp in length
* 3644 3743: gap of 100 bp
* 3744 6011: contig of 2268 bp in length
* 6012 6112: gap of 100 bp
* 6112 7495: contig of 1384 bp in length
* 7496 7595: gap of 100 bp
* 7596 10610: contig of 3015 bp in length
* 10611 10710: gap of 100 bp
* 10711 15590: contig of 4880 bp in length
* 15591 15690: gap of 100 bp
* 15691 22192: contig of 6502 bp in length
* 22193 22292: gap of 100 bp
* 22293 48685: contig of 26393 bp in length
* 48686 48785: gap of 100 bp
* 48786 87575: contig of 38790 bp in length
* 87576 87676: gap of 100 bp
* 87676 133888: contig of 46213 bp in length.

FEATURES

Location/Qualifiers

source

1. .133888

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="15"

/map="15"

/clone="RP11-28B17"

/clone_1lb="RPC1-11 Human Male BAC"

1. .1054

/note="assembly_fragment"

1155. .2415

/note="assembly_fragment"

2516. .3643

/note="assembly_fragment"

3744. .6011

/note="assembly_fragment"

6112. .7495

/note="assembly_fragment"

7596. .10610

/note="assembly_fragment"

10711. .15590

/note="assembly_fragment"

clone_end:SP6

vector_side:left"

15691. .22192

/note="assembly_fragment"

22293. .48685

/note="assembly_fragment"

48786. .87575

/note="assembly_fragment"

clone_end:T7

vector_side:left"

87676. .133888

/note="assembly_fragment"

ORIGIN

Query Match 54.3%; Score 1260; DB 2; Length 133888;

Best Local Similarity 100.0%; Pred. No. 1.3e-253;

Matches 1260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 316 TTATATTTTGTAGATGCTTTCATAGAAAGCTTCGAATCAACGGAAGTGGATTAATAG 375
Db 2414 TTATATTTTGTAGATGCTTTCATAGAAAGCTTCGAATCAACGGAAGTGGATTAATAG 2355
QY 376 TGGTGATCCAGAAATTTGTCACCTCAAGTAATAAACAACCTCAGAGTTAAACACCTAAG 435
Db 2354 TGGTGATCCAGAAATTTGTCACCTCAAGTAATAAACAACCTCAGAGTTAAACACCTAAG 2295
QY 436 TTCTCACCAGATGTCATATATCCGACGATTAATTTATGCTTCTTGACCGTAGG 495
Db 2294 TTCTCACCAGATGTCATATATCCGACGATTAATTTATGCTTCTTGACCGTAGG 2235
QY 496 ACTTCCACATGACAGATTTTGAACCTCTTGACGTAAGTGGGGAATGAGTGCATTT 555
Db 2234 ACTTCCACATGACAGATTTTGAACCTCTTGACGTAAGTGGGGAATGAGTGCATTT 2175
QY 556 ATTGCTACACCATGCGTCATCGAGTAAGTGGCTTAATGTTGCTAAAAAAGCAGAGAGC 615
Db 2174 ATTGCTACACCATGCGTCATCGAGTAAGTGGCTTAATGTTGCTAAAAAAGCAGAGAGC 2115
QY 616 AAAAGTGATGACATGAACCTCTGTTTCATCGAATAAATAATGACTTTTCCGCAAG 675
Db 2114 AAAAGTGATGACATGAACCTCTGTTTCATCGAATAAATAATGACTTTTCCGCAAG 2055
QY 676 ACATCTACTAATCTTGATAGTAATAAACAATTGTCATGTCAGAGGCTCATTCGAGAGC 735
Db 2054 ACATCTACTAATCTTGATAGTAATAAACAATTGTCATGTCAGAGGCTCATTCGAGAGC 1995
QY 736 AGATATTTCTGGATATGTCATAAACAATTAGTAATTATCATCGTCCACTGAATC 795
Db 1994 AGATATTTCTGGATATGTCATAAACAATTAGTAATTATCATCGTCCACTGAATC 1935
QY 796 TGTGTTCAATACGTTAACTCTTCATATTAGAAATGAGGCTGATGAGTTCATATTTT 855

Db	1934	TGTGGTTCATTACGCTCTTAACCTTCATATTAGAAATGAGGCTGATGAGTCCATATTT	1875
QY	856	GAAAGTTTTCATCAGTACTAGTTTTTTGATAGCTCAAGCCAGAGTGTCTTTCTA	915
Db	1874	GAAAGTTTTCATCAGTACTAGTTTTTTGATAGCTCAAGCCAGAGTGTCTTTCTA	1815
QY	916	TCTACTCTCATACACCAATAAATGCTGAATGAATCTAAGCGGAGATCGCCTAGTAT	975
Db	1814	TCTACTCTCATACACCAATAAATGCTGAATGAATCTAAGCGGAGATCGCCTAGTAT	1755
QY	976	TTTAACTATTGCTGGCAGCATTTCTGAGTCCAATATATAAGTATGTGTACCTTTGCT	1035
Db	1754	TTTAACTATTGCTGGCAGCATTTCTGAGTCCAATATATAAGTATGTGTACCTTTGCT	1695
QY	1036	GGGTCAAGTTGTTCTTTAGAGAGAGTAAAGGATCAAAATGCACTAAACGAACTGAACA	1095
Db	1694	GGGTCAAGTTGTTCTTTAGAGAGAGTAAAGGATCAAAATGCACTAAACGAACTGAACA	1635
QY	1096	AGCGATCGAAATATCCCTTTGGGATTTCTTGACTCGATAGTCTATTATTTTCAGAGAA	1155
Db	1634	AGCGATCGAAATATCCCTTTGGGATTTCTTGACTCGATAGTCTATTATTTTCAGAGAA	1575
QY	1156	AAATATTCATTGTTTCTGGGTTGGTGAATTCACCAATCATTCATTCAAAAATTGTTGT	1215
Db	1574	AAATATTCATTGTTTCTGGGTTGGTGAATTCACCAATCATTCATTCAAAAATTGTTGT	1515
QY	1216	TTACCAACACCCATTCGCCCGCGATAAAGCATGAATGTTCCGTCTGGGCATAGATTAAC	1275
Db	1514	TTACCAACACCCATTCGCCCGCGATAAAGCATGAATGTTCCGTCTGGGCATAGATTAAC	1455
QY	1276	GTCACCTCAAAAAGGTATAGTTAAATCACTGAATCCGGGAGCACTTTTCTATTAATGAA	1335
Db	1454	GTCACCTCAAAAAGGTATAGTTAAATCACTGAATCCGGGAGCACTTTTCTATTAATGAA	1395
QY	1336	AAGTGGAAATCTGACAAATCTGGCAAAACCATTTTAACACACGTGCGAACTGTCATGAAT	1395
Db	1394	AAGTGGAAATCTGACAAATCTGGCAAAACCATTTTAACACACGTGCGAACTGTCATGAAT	1335
QY	1396	TCTGAAAGAGTTACCCCTCTAAGTAATGAGGTGTTAAGACGCTTTCATTTTCAATGTCG	1455
Db	1334	TCTGAAAGAGTTACCCCTCTAAGTAATGAGGTGTTAAGACGCTTTCATTTTCAATGTCG	1275
QY	1456	GCTAATCGATTTGGCCATACTAATAATCCTGAATAGCTTTAAGAAGTTATGTTAAAA	1515
Db	1274	GCTAATCGATTTGGCCATACTAATAATCCTGAATAGCTTTAAGAAGTTATGTTAAAA	1215
QY	1516	CCATCGCTTAATTGCTGAGATTAAACATAGTAGTCAATGCTTTCACTAAGGAAAAAAC	1575
Db	1214	CCATCGCTTAATTGCTGAGATTAAACATAGTAGTCAATGCTTTCACTAAGGAAAAAAC	1155

RESULT 5

CVU39284

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCES

AUTHORS

TITLE

42529 bp DNA linear SYN 13-APR-1996

Cloning vector TLF97-1, lambda phage lacZ translational fusion vector, complete sequence.

U39284

U39284.1 GI:1066304

Cloning vector TLF97-1

Cloning vector TLF97-1

artificial sequences; vectors.

1 (bases 1 to 42529)

St Pierre,R. and Linm,T.

A refined vector system for the in vitro construction of single-copy transcriptional or translational fusions to lacZ

Gene 169 (1), 65-68 (1996)

8635751

2 (bases 1 to 42529)

StPierre,R.

Direct Submission

JOURNAL	Submitted (24-OCT-1995) Thomas Linm, Microbiology and Immunology, University of Western Ontario, London, Ontario N6A5C1, Canada
FEATURES	
source	1. .42529
terminator	/organism="Cloning vector TLF97-1"
gene	/mol_type="genomic DNA"
	/db_xref="taxon:43840"
	complement(19404..19489)
	/note="rnbT2 terminator"
	complement(19507..20760)
	/gene="lacY"
	complement(19507..20760)
	/gene="lacY"
	/codon_start=1
	/transl_table=11
	/product="lactose permease"
	/protein_id="AAC53648.1"
	/db_xref="GI:1066305"
	/translation="MYLLKNTNFMMFGLFEFFVFIMGAYFPFPIMLHDINHISKSD
	TGIIFAAISLSLFLQPLFGLSDKLGKYLMIITGMLVMPAPFPIIFGDLIQYN
	ILVGSIVGGIYLGFCENAGAPAVEAFIEKYSRBSNFEGRAMFGVGNALCASIVGI
	MFTINNOFVFMGSCALILAVLLFPAKTDAPSSATVANAVGANSHAFSLKALIELFR
	QPKLWPLSLYIVGSCITYDVFDDQFANFTSFATGEQGTREVEGYVTMGELLNASIM
	FFAPLIIINRIGKNALLLAGTIMSLIGSSFATSALAEVILKTIHMEFVEPDLVGC
	KYITSQFEVRSATILYVCFCEFKQLMIFMSVLAGNMYESIGFQAYLVGLVALGF
	TLISVFTLGSGPLSLRQVNEVA"
	complement(20812..23859)
	/gene="lacZ"
	complement(20812..23859)
	/gene="lacZ"
	/note="EcoRI site normally present at the 3' end of lacZ
	has been removed by an uncharacterized mutation"
	/codon_start=1
	/transl_table=11
	/product="beta-galactosidase"
	/protein_id="AAC53649.1"
	/db_xref="GI:1066306"
	/translation="VVLQRRDWEENPGVTQLNRLAAHPPEFASWRNSEARDRPSQQLR
	SLNGEWFAMFPAPBAEVESWLECDLPEADTVVVPSPNMOMHGYPADPIYNNVYPIITVN
	PPEVPTENPTGCSLTFFNVDESWLOEQTRIIFDGVNSAFHLWCNGRWYGGQDSRLP
	SEFDLSAFLRAGENRLAVMLRMSDGSYLEDDQDMRMSGIFRDVSLHKPTQISDFH
	VATRFNDPFSRAVLEAEVOMCGELRDYLRVTVSLQGETOVAAPFGGELIDERGG
	YADRVTLRLENVPKLMSEAIENLYRAVLELHTADGTLIEAECADVGFREYRIENGILL
	LINGKPLIRGVNRHEHPILHGQVMEQTMVODILMKONNENAVRCSHYPHMLMYT
	LCDRYGLYVVDENANIEHGMVPMNRLTDPRMLPAMSERVTRMVQRDRNPSYIIMSL
	GNESGHANBDALYRWIKSVDPSPVQYEGGADTTATDIIICPMYARVDEDDPPAPAV
	KMSIKKWLILPGETRPLILCEYAHAMGNSLGGFAKYQAFRQYPRLOGGFPWIDVDS
	LIKYDENGPNWSAYGDFEGDTPNDROFCMNGLVFADRTPHPALTEAKHQOQFRQRLS
	GQTIIEVTSLEYLFRHSDELHLMWVALDKPLASGEVPLDVAPQKQILIEPLDPES
	AGQLMTVRVVOGNATAMSEBAGHISAWQOMRLAENLSVTLPAASHAIPLHLSMDFC
	IELGNKRWQFNROSGFLQMWITDCKQLTPRPDQFTRPAQHLNDIGVSEATRIDENAM
	VERWKAAGHYQAEALLQCTADTLADAVLITTAHAMQHQKTLFISRKTYRIDSGQM
	AITVDVEVASDTPHPARIGLNCQAOVAERVNWLGLQENYPDRLTAACFPWDLP
	SDMYTPYVFPSENGIRCGTRELINYPHQWRGDFQFNISRSQQQLMETSHRHLHAE
	GTWLNIDGFHMGIQGDSDWSPSVSAEFQLSAGRXYQLVWCQK"
	complement(23874..23957)
	/note="multiple cloning site"
	complement(23973..24141)
	/note="rnbT1 terminator"
	24475..25665
	/gene="tetR"
	24475..25665
	/gene="tetR"
	/codon_start=1
	/transl_table=11
	/product="tetracycline resistance protein"
	/protein_id="AAC53650.1"
	/db_xref="GI:1066307"
	/translation="MKSNNALIVILGTVTLDAVGIGLVMPVLPGLLRDIYHSDSIASH
	YGVLLATYALMQFLCAPVLGALSDFRGRPVLLASLGLATIDYAIMATTPVLMILVAG
	RIVAGITGATGAVAGAYIADITDGEDRARHFGMSACFGVGVAVGAGLGAISLH
	APFLAAVLNGLNLILGCFLMQESHKGERRPMPLRAFPVSSFRWARGMTIVAALMTV
	FFIMQLVGVPAALMWIFGEDRFRWSATMIGLSLAVRGILHAAQAQAVTGPATKRFGE

KOALIIAGMADALGYVLLAFATRGMMAPFIMILLASGIGMPALQAMLSHOVDDHOG
QLOGSLAALTSLSITIGPLIVTAIYAASASTWNGIAMIYGAALYLVCLPALIRGAMSR
ATST"

misc_feature 33602..35450
/note="imm21 substitution"

ORIGIN

Query Match 40.8%; Score 947; DB 12; Length 42529;
Best Local Similarity 100.0%; Pred. No. 4.9e-188;
Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1376 GTGCGAACTGTCCATGAATTTCTGAAAAGAGTTACCCCTTAAGTAATGAGGTGTAAGA 1435
|||||
Db 25763 GTGCGAACTGTCCATGAATTTCTGAAAAGAGTTACCCCTTAAGTAATGAGGTGTAAGA 25822
QY 1436 CGCTTTCAATTTCAATGTCGGGTAATCGATTGGCCACTACTAATCCTGAATGCTT 1495
|||||
Db 25823 CGCTTTCAATTTCAATGTCGGGTAATCGATTGGCCACTACTAATCCTGAATGCTT 25882
QY 1496 TAAGAAAGTTATGTTAAACCAATCGCTTAATTGCTGAGATTAACTAGTAGTCAATGC 1555
|||||
Db 25883 TAAGAAAGTTATGTTAAACCAATCGCTTAATTGCTGAGATTAACTAGTAGTCAATGC 25942
QY 1556 TTTCACCTAAGAAAAAAACATTTCAGGAGTTGACTGAATTTTATCTAATTAAGAAT 1615
|||||
Db 25943 TTTCACCTAAGAAAAAAACATTTCAGGAGTTGACTGAATTTTATCTAATTAAGAAT 26002
QY 1616 AAGTGCTTACTTCTCTTTTGAACCTACAAAACCAATTTTAAACATTTCCGATATCGCAT 1675
|||||
Db 26003 AAGTGCTTACTTCTCTTTTGAACCTACAAAACCAATTTTAAACATTTCCGATATCGCAT 26062
QY 1676 TTTCACCATGCTCATCAAGAAGAGTAAGATAAAACATTTGTAACAAGAATAGTCAATCC 1735
|||||
Db 26063 TTTCACCATGCTCATCAAGAAGAGTAAGATAAAACATTTGTAACAAGAATAGTCAATCC 26122
QY 1736 AACCATCTGCTCGTAGGAATGCTTATTTTCTACTGAGGAATATACCCGCTCTTT 1795
|||||
Db 26123 AACCATCTGCTCGTAGGAATGCTTATTTTCTACTGAGGAATATACCCGCTCTTT 26182
QY 1796 CAATTAACACTAACTCCAACTATAGTAACCCCTTAATTTTAAATTAACCGCAATTTA 1855
|||||
Db 26183 CAATTAACACTAACTCCAACTATAGTAACCCCTTAATTTTAAATTAACCGCAATTTA 26242
QY 1856 TTGGCGGCAACACAGATCTCTTTTAAAGTACTCTCTATTAACATACGTTTCCATCT 1915
|||||
Db 26243 TTGGCGGCAACACAGATCTCTTTTAAAGTACTCTCTATTAACATACGTTTCCATCT 26302
QY 1916 AAAAATTAGTAGTATTGAAGCTTAACGGGCGATCGTATGTAATTTCCATATTTAGCTTT 1975
|||||
Db 26303 AAAAATTAGTAGTATTGAAGCTTAACGGGCGATCGTATGTAATTTCCATATTTAGCTTT 26362
QY 1976 CTGCTTCTTTTGATTAACCACTGTTATTCATGTTGATGTTGCACTGTTTATTAACCAAC 2035
|||||
Db 26363 CTGCTTCTTTTGATTAACCACTGTTATTCATGTTGATGTTGCACTGTTTATTAACCAAC 26422
QY 2036 GATATAGTCTATTAATGATATATAGTATCGCCGAACGATTAGCTCTTCAAGGCTTCTGAA 2095
|||||
Db 26423 GATATAGTCTATTAATGATATATAGTATCGCCGAACGATTAGCTCTTCAAGGCTTCTGAA 26482
QY 2096 GAAGCGTTTCAAGTACTAAATAAGCCGATAGATAGCCAGGACTTCTGAGCCATTTTTCAT 2155
|||||
Db 26483 GAAGCGTTTCAAGTACTAAATAAGCCGATAGATAGCCAGGACTTCTGAGCCATTTTTCAT 26542
QY 2156 AAGTGTTAATCTCCGCTCTGCTCATTAACAGACATTCATCACTAGTTATGCGGAAAGGT 2215
|||||
Db 26543 AAGTGTTAATCTCCGCTCTGCTCATTAACAGACATTCATCACTAGTTATGCGGAAAGGT 26602
QY 2216 ATGCATGCTGGGTGTGGGGAAGTCTGGAAGAAAAGAGTCAAGTGCCTGCTTTGACATC 2275
|||||
Db 26603 ATGCATGCTGGGTGTGGGGAAGTCTGGAAGAAAAGAGTCAAGTGCCTGCTTTGACATC 26662
QY 2276 ACTGCTATCTTCTTACTGTTATGCAAGGTCGTAGTGGGTGGCACACA 2322
|||||

Db 26663 ACTGCTATCTTCTTACTGTTATGCAAGTCTGAGTGGGTGGCACACA 26709

RESULT 6
CVU39285
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCES
AUTHORS
TITLE
JOURNAL
FEATURES
Source
terminator
gene
CDS
gene
CDS

CVU39285
Cloning vector TLF97-2, phage lambda lacZ translational fusion
vector, complete sequence.
U39285
U39285.1 GI:1066308
Cloning vector TLF97-2
Cloning vector TLF97-2
artificial sequences; vectors.
1 (bases 1 to 42530)
St Pierre, R. and Linn, T.
A refined vector system for the in vitro construction of
single-copy transcriptional or translational fusions to lacZ
Gene 169 (1), 65-68 (1996)
96186904
8635751
2 (bases 1 to 42530)
St Pierre, R.
Direct Submission
Submitted (24-OCT-1995) Thomas Linn, Microbiology and Immunology,
University of Western Ontario, London, Ontario N6A5C1, Canada
location/Qualifiers
1. 42530
/organism="Cloning vector TLF97-2"
/mol_type="genomic DNA"
/db_xref="taxon:43841"
complement(19404..19489)
/note="xrbT2 terminator"
complement(19507..20760)
/gene="lacy"
complement(19507..20760)
/gene="lacy"
/codon_start=1
/transl_table=11
/product="lactose permease"
/protein_id="AAC53651.1"
/db_xref="GI:1066309"
/translation="MYILKNTNFMFGLFEFFYFIMGAYFPFPIWLDINHISKSD
TGLIFAIISFSLFQPLFGLSDKGLRKYLIIITGMVMAFPFIIFGLLOYN
ILVGSIVGGIYLGFCFNAGAPAVEAFIEKVSRSNFEFGARMFGCVWALCASIVGI
MFTINNOFVFMISGSCALLIAYLIFRAKIDAPSSATVANAVGANHSAFSLKLELFR
QPLWFLSLYIVGSCYDVFDOQFANFTSFATGEOGTRVFGYVTMGEILLNASIM
FAPLIINRIGGKNALLIAGTIMSVRIIGSSFATSALEVLIKTLHMEVFPFLVGCF
KYTSQFEVFPSATIYLVCFEKFQKLMIFMSVLAGNMYESIGFGAYLVGLVALGF
TLISVFTLSGPGPLSLRRQVNEVA"
complement(20812..23859)
complement(20812..23859)
/gene="lacz"
complement(20812..23859)
/gene="lacz"
/note="EcoRI site normally present at the 3' end of lacZ
has been removed by an uncharacterized mutation"
/codon_start=1
/transl_table=11
/product="beta-galactosidase"
/db_xref="GI:1066310"
/translation="VVLQRDWMENPGVTQNLRLAAHPFASWRNSEARDRPSQOLR
SLNGEMRFAPFAPEAVPESWLECDLPEADIVVPSNMOMHGIDAPITVNTYPTVN
PPFVPTENPGCYSLTFNVDSEWLOEGGRIIFDGVNSAFHLMCNGRWYGGQDSRLP
SEFDLSAFLEAENRLEAVMLRWSDGSYLEDQDMRMSGRTVRSALPSTVILMSL
VATRENDPFRALVLAEVQMGELRDYLRVTVSLMOGETQVASTGAPGGEIIDERG
YADRVTLRLVNEPKLMSAEIPNLRYRAVELHTADGTLIEAECADVREVRILEGLL
LLNGKPLLRGNRHEHPLHGQVMEQTVODILMKONNENAVRCSHYENHPLMYT
LCDRYGLYVDEANIEHGMVPMNRLTDERMLPAMSERVTRMVOBRNHSVLIWSL
GNESGHGANHDALYRIKSVDPSPVQYEGGADTTATDIIICMYARVDEDQPPPAV
KMSIKKWLSPGETRPLILCEYAHAMGNSLGGFAKYWOAFROYPRLOGGFVMDVDS
LIKDYDENGNSAYGGDFGDTPNDRQFCMNGLVFADRTPHALTEAKHQOQFFQFRLS
GQIEVTSEYLFERHSDNELHMMVVALDGKPLASGEVPLDVAPOGKLELLELPQES


```

gene
  TLISVFTLSGPGPLSLRRQVNEVA"
  complement (20812. .23859)
  /gene="lacZ"
  complement (20812. .>23859)
  /gene="lacZ"
  /note="EcoRI site normally present at the 3' end of lacZ
  has been removed by an uncharacterized mutation"
  /codon_start=1
  /transl_table=11
  /product="beta-galactosidase"
  /protein_id="AAC53655.1"
  /db_xref="GI:1066314"
  /translation="VVLQRRDMENPGVTQLNRLAHPPEFASWRNSEARTDRPSQQLR
  SLNGEMRFAMFPAPAEVBSWLECDLPEADTVVPSNMQMHGYDAPITYNTYPTVN
  PEPVPTENPGCYSLFENVDESMLQEGQTRIFDGVNSAFHLMCNGRWVGQDSRLP
  SEFDLSAFLRAGENRLAVMLRMSDSYLEDQMRMSGI FRDVSLLHKPTQISDFH
  VATRFNDFFSRAVLEAEVQMGELRDYLRVTSIMQGETOASGTAPFGELIDERG
  YADRVTLRLNVENPKLMSAEIPNLRYRAVELHTADGTLIEAACDVGFREYRIENGL
  LNLGKPLLRGVNRHSHPLHGQVMDQTMVODILLMKONNFNAVRCSHYDNHPLYMT
  LCDRYGLYVVDENAIETHGMVPMNRLTDDPRWLPAMSERVTRMVRQDRNPSVITWSP
  GNESGHGANHDALYRWIKSVDPSPVYEGGADTTATDII CPMYARVDEHQPEPAV
  KWSIKKMLSLPGETRPILICEYAHAMGNSLGFAKYQAFROY PRLQGFVWDVDS
  LKIDENGNPWSAYGGDFGDTPNDRQFCMNGLYFADRTPHALTEAKHQQFFQFRLS
  GQTEVTSXYLFRHSDELHMMVALDGKPLASGEVPLDVAPOGKQLELDELPOES
  AGQLMTVRVQPNATAMSEAGHISAMQWRLAENLSVTLPAASHAIPLHTSEMDFC
  IELGNKRWQENKQSGFLSQMWIGIDSKQLTPLRQFTRAPLNDIGVSEATRIDPNAM
  VERMKAAGHYQAEAAALLOCTADTLADAVLITTAAMQHQGKTLFISRTKYRIDSGQM
  AITVDVEVASDTPHARIGLNCOLAQVAERVWGLGPOENY PDLRTAACFDRLPL
  SDMTYPRVPESENGLRCGTRELNPGHQWRGDFQFNISRYSQQQLMETSHRHLHAEE
  GTWLNIDGFHMGIGGDDSWSPSVSAEFQLSAGRYHYQLVWCQK"
  complement (23876. .23959)
  /note="multiple cloning site"
  complement (23975. .24143)
  /note="rrnB1 terminator"
  24477. .25667
  /gene="tetR"
  24477. .25667
  /gene="tetR"
  /codon_start=1
  /transl_table=11
  /product="tetracycline resistance protein"
  /protein_id="AAC53656.1"
  /db_xref="GI:1066315"
  /translation="MKSNNALIVILGVTLLDAVGIGLWMPVL PGLLRDIVHSDSIASH
  XGVLLAYALMQFLCAPVLGALSDRFGRPVLLASLLGATIDYAIMATPEVLMILYAG
  RIVAGITGATGAVAGAYIADITDGEDRARHFGMSACFGVGVAVGAVGGLGALSLH
  APFLAAVNLGNLNLGCFLMQESHKGERRPMPLRAFNPVSSFRMARGMTIVAALMTV
  EFIMQLVGVPAPALMWIFEGEDRFWSATMIGSLAVFGILHALAQAFTGPKRFGF
  KOAIIAGMAADALGYVLLAFATRGWMAFPIMILLASGGIGMPALQAMISROVDDHQG
  QLOGSLAALTSLTSITGPLIVTAIYAASASTWNGLAMI VGAALYLVCLPALRGAMSR
  ATST"
  33604. .35452
  /note="imm21 substitution"

misc_feature
  33604. .35452
  /note="imm21 substitution"

ORIGIN
Query Match      40.8%; Score 947; DB 12; Length 42531;
Best Local Similarity 100.0%; Pred.No. 4.9e-188;
Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1376 GTGCGAACTGTCATGAATTTCTGAAAGAGTTACCCCTCTAAGTAATGAGGTGTTAAGA 1435
      |||||||
Db 25765 GTGCGAACTGTCATGAATTTCTGAAAGAGTTACCCCTCTAAGTAATGAGGTGTTAAGA 25824

QY 1436 CGCTTTCATTTCAATGTCGGCTAATCGATTGGCCATACTAAATCCGGAATAGCTT 1495
      |||||||
Db 25825 CGCTTTCATTTCAATGTCGGCTAATCGATTGGCCATACTAAATCCGGAATAGCTT 25884

QY 1496 TAAGAAAGTTATGTTAAACCATCGCTTAATTTGCTGAGATTAAACATAGTAGTCATGC 1555
      |||||||
Db 25885 TAAGAAAGTTATGTTAAACCATCGCTTAATTTGCTGAGATTAAACATAGTAGTCATGC 25944

QY 1556 TTTCACCTAAGAAAAAACATTTCAGGAGTGCATGAATTTTATCTATTAATGAAT 1615
      |||||||
```

```

Db 25945 TTTCACCTAAGAAAAAACATTTCAGGAGTGCATGAATTTTATCTATTAATGAAT 26004
QY 1616 AAGTGCTTACTTCTTTTGGACCTACAAACCAATTTTAACATTTCCGATATGCAAT 1675
      |||||||
Db 26005 AAGTGCTTACTTCTTTTGGACCTACAAACCAATTTTAACATTTCCGATATGCAAT 26064
QY 1676 TTTCACCATGCTCATCAAGACAGTAAGATAAAACATTTGTAACAAAGGAATAGTCATCC 1735
      |||||||
Db 26065 TTTCACCATGCTCATCAAGACAGTAAGATAAAACATTTGTAACAAAGGAATAGTCATCC 26124
QY 1736 AACCATCTGCTCGTAGGAATGCCCTTATTTTCTACTGCAAGAAATATACCCGCCCTCTT 1795
      |||||||
Db 26125 AACCATCTGCTCGTAGGAATGCCCTTATTTTCTACTGCAAGAAATATACCCGCCCTCTT 26184
QY 1796 CAATAACACTAACTCCACATATAGTAACCCCTTAATTTTAAATAACCGCAATTTA 1855
      |||||||
Db 26185 CAATAACACTAACTCCACATATAGTAACCCCTTAATTTTAAATAACCGCAATTTA 26244
QY 1856 TTGGCGGCAACACAGAGATCTCTTTTAAGTTACTCTCTATTAACATACGTTTCCATCT 1915
      |||||||
Db 26245 TTGGCGGCAACACAGAGATCTCTTTTAAGTTACTCTCTATTAACATACGTTTCCATCT 26304
QY 1916 AAAAATTAGTAGTATTGAACTTAACCGGGCATCGTATTGTAAGTTTCCATATTAGCTTT 1975
      |||||||
Db 26305 AAAAATTAGTAGTATTGAACTTAACCGGGCATCGTATTGTAAGTTTCCATATTAGCTTT 26364
QY 1976 CTGCTTCCTTTTGATAACCCCATGTTATTCAATGTTGATGCTGACATGTTTATACCAAC 2035
      |||||||
Db 26365 CTGCTTCCTTTTGATAACCCCATGTTATTCAATGTTGATGCTGACATGTTTATACCAAC 26424
QY 2036 GATATAGTCTATTAATGATATATAGTATCGCCGAACGATTAAGCTCTTCAGGCTTCTGAA 2095
      |||||||
Db 26425 GATATAGTCTATTAATGATATATAGTATCGCCGAACGATTAAGCTCTTCAGGCTTCTGAA 26484
QY 2096 GAAGCGTTTCAAGTACTAATAAGCCGATAGATAGCCACGGACTTCGTAGCCATTTTCAT 2155
      |||||||
Db 26485 GAAGCGTTTCAAGTACTAATAAGCCGATAGATAGCCACGGACTTCGTAGCCATTTTCAT 26544
QY 2156 AAGTGTTAAGTTCGCTCGCTCGCTCATACAGACATTTCACTACAGTTATGGCGGAAGGT 2215
      |||||||
Db 26545 AAGTGTTAAGTTCGCTCGCTCGCTCATACAGACATTTCACTACAGTTATGGCGGAAGGT 26604
QY 2216 ATGCATGCTGGGTGTGGGGAAGTCGTGAAAGAAAAGAGTCAGCTGCGTCTTGGACATC 2275
      |||||||
Db 26605 ATGCATGCTGGGTGTGGGGAAGTCGTGAAAGAAAAGAGTCAGCTGCGTCTTGGACATC 26664
QY 2276 ACTGCTATCTTCTTACTGCTTATGACAGGTGCTGATGGGTGGCACACA 2322
      |||||||
Db 26665 ACTGCTATCTTCTTACTGCTTATGACAGGTGCTGATGGGTGGCACACA 26711

RESULT 8
CVU37692
LOCUS CVU37692 42704 bp DNA linear SYN 13-APR-1996
DEFINITION Cloning vector lambda TXF97, lacZ transcriptional fusion vector,
complete sequence.
ACCESSION U37692
VERSION U37692.1 GI:1051181
KEYWORDS Cloning vector lambda TXF97
SOURCE Cloning vector lambda TXF97
ORGANISM Cloning vector lambda TXF97
          artificial sequences; vectors.
REFERENCE 1 (bases 1 to 42704)
AUTHORS St Pierre, R. and Linn, T.
TITLE A refined vector system for the in vitro construction of
single-copy transcriptional or translational fusions to lacZ
JOURNAL Gene 169 (1), 65-68 (1996)
MEDLINE 96186904
PUBMED 8635751
REFERENCE 2 (bases 1 to 42704)
AUTHORS St Pierre, R. and Linn, T.
TITLE Direct Submission
JOURNAL Submitted (29-SEP-1995) Thomas Linn, Microbiology and Immunology,
```

University of Western Ontario, London, Ontario N6A5C1, Canada
Location/Qualifiers
1. .42704
/organism="Cloning vector lambda TXF97"
/mol_type="genomic DNA"
/db_xref="taxon:43300"
1. .42704
/note="EcoRI site normally present at this site has been removed by an uncharacterized mutation"
misc_feature
5505
/note="BamHI site normally present at this site has been removed by an uncharacterized mutation"
terminator
complement(19404. .19489)
/note="rlnBT2 terminator"
complement(19507. .20760)
/gene="lacZ"
complement(19507. .20760)
/gene="lacY"
/codon_start=1
/transl_table=11
/product="lactose permease"
/protein_id="AAC53645.1"
/db_xref="GI:1051182"
/translation="MYLLKNTNFMFGLFFFEYFIMGAYFPFPFIMLHINHSKSD
TGIIFAAISLFLFQPLFGLDKLGLRXYLMIITGLVMFAPFIFLFGPLQYN
ILVGSIVGIGYILGFCFNAGAPAVEAFIEKVSRSNFEFGRARFGCVGNALCASIVGI
MFTINNOVFVFWLSSGICALILAVLFFAKTDAPSSATVANAVGANHSAPSLALELFR
QPKMFLSLVIVGSCYDVDPDQOFANFPTSFPATGEQGTGVFGYVTMGLLNASIM
FEAPLIINRIGGNALLAGTISSVRIIGSSFATSALFVILKTIHMFEPFLVGC
KYITSQFEVRSATIIYLVCFEFGKQLAMIFMSVILAGNMYESIGFGAYLVGLVALGF
TLISVFTLSGPGPLSLRQVNEVA"
complement(20812. .23886)
/gene="lacZ"
complement(20812. .23886)
/gene="lacZ"
/note="EcoRI site normally present at the end of lacZ has been removed by an uncharacterized mutation"
/codon_start=1
/transl_table=11
/product="beta-galactosidase"
/protein_id="AAC53646.1"
/db_xref="GI:1051183"
/translation="MTWITDSLAVLQDRDMEVGTQLNRLAHPPEASWENSEBAR
TDRPSQRLSINGENRFAMFAPPAVEPESWLECDLPEADTVVPSNMQMHGDAPIYT
NVTYPITVNPFPVPTENPTGCYSLTFNDESWLOEQTRIIFDGVNSAFHMCNGRW
GYGQDSRLPSEFDSAFIRAGENRLAVMLRMSDGSYLEDDQMMRMSGIRPDVSLHK
PTTQISDFHVAITRENDPFSRAVLEAEVOMCGELRDYLRTVSLMOGETOYASGTA
GEIIDERGGYADRVTLRLNVENPKMSAEIPNLRYRAVELHTADGTLIEAACDVGFR
EVRINGLLNGKPLIRGVNRHHPHPLHGVDQTMODILMKONNENAVRCSH
YVNPPLWYTLCDRYGLVYVDEANIEHGMVPMNRLLTDEPRWLPAMSERVTRMVQDRN
HPSVITMSLGNESGHANHDALYRMIKSVDPSPVQYEGGADTTATDITCPMYARVD
EDQFPVAPPKMSIKKWLSPGETRPLILCEYAHAMGNSLGGFAKYQWAFROYPRLOGG
FVMDWVDOSLILKYDENGPNPWSAGYDGFDTPNDRQFCMNGLVFADRTPHPLTEAKHO
QQFFQRLSGQTIETVTSYLLFRHSDNELHMVVALDGKPLASGEVPLDVADQKOLIE
LPELPQESAGOLMLTVRVVQPNATAMSEAGHISAMQWRILAENLSVTLPASHALPH
LTTSEMDFCIELGNKRWQFNROSGFLSOMWIGDKQLLTPLRDQFTRAPLNDIGVSE
ATRIDPNAMVERMKAGHYQAEALLOCTADTLADAVLITTAHAMQHGKTLFISRKT
YRIDSGQMAITVVEVASDTPHPARIGLNCOLAQVAERVNMLGLPOENYPDRLTAA
CFDRWDLPLSDMYTPYVFPSENGLRCGTRELANYGPHQWRGDFQFNISRYSOOLMETS
HRHLHAEBEGTWNLDGFHMGIGGDDSWSPSVSAEFQLSAGRHYQLVWCQK"
complement(24049. .24132)
/note="multiple cloning site"
complement(24148. .24316)
/note="rlnBT1 terminator"
24650. .25840
/gene="tetR"
24650. .25840
/gene="tetR"
/note="Description: tetracycline resistance gene;
tetracycline resistance protein"
/codon_start=1
/transl_table=11
/protein_id="AAC53647.1"

/db_xref="GI:1051184"
/translation="MKSNNALIVILGVTTLDAVIGLWMPVLPGLRDIIVHSDSIASH
YGVLLALYALMQFLCAPVLGALSDRFRREPVLLASLGATIDYALNATTPVLMILYAG
RIVAGITGATGAVAGAYIADITDGEDRARFGLMSACFVGWVAGPYAGGILGSLH
APFLAAVLNGLNILLGCFPLMOESHKGERPMDLRAFNPVSSFRWARGMTIVAALMTV
FFIMQVLGVPPALMWIFGEDRFRMSATWIGSLAVFGIHALAQAFYGPATKRFGE
KQAIAGMADALGVLLAFATRGWMAFPIMILLASGGIGMPALQAMSRQVDDHOG
QLQGSIALTSLSITGPLIVTAIYAAASASTWNGLAWIVGALYLVCLPALRGAMSR
ATST"
misc_feature
31747
/note="EcoRI site normally present at this site has been removed by an uncharacterized mutation"
misc_feature
33777. .35625
/note="imm21 substitution"
39168
/note="EcoRI site normally present at this site has been removed by an uncharacterized mutation"
ORIGIN
Query Match 40.8%; Score 947; DB 12; Length 42704;
Best Local Similarity 100.0%; Pred. No. 4.9e-188;
Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1376 GTGCGAACTGTCATGAATTTCTGAAAGAGTTACCCCTCTAAGTAATGAGGTGTTAAGGA 1435
|||||
Db 25938 GTGCGAACTGTCATGAATTTCTGAAAGAGTTACCCCTCTAAGTAATGAGGTGTTAAGGA 25997
QY 1436 CGCTTTCATTTCAATGTCGGCTAATCGATTTGGCCATACTAATAATCCTGAATAGCTT 1495
Db 25998 CGCTTTCATTTCAATGTCGGCTAATCGATTTGGCCATACTAATAATCCTGAATAGCTT 26057
QY 1496 TAAGAAGGTATGTTTAAACCATCGCTTAATTTGCTGAGATTAAACATAGTAGTCATGC 1555
Db 26058 TAAGAAGGTATGTTTAAACCATCGCTTAATTTGCTGAGATTAAACATAGTAGTCATGC 26117
QY 1556 TTTCAACCTAAGAAAAAACATTTCAAGGAGTGTACTGAATTTTATCTATTAATGAAT 1615
|||||
Db 26118 TTTCAACCTAAGAAAAAACATTTCAAGGAGTGTACTGAATTTTATCTATTAATGAAT 26177
QY 1616 AAGTGCTTACTTCTTTTGTGACCTACAAACCAATTTTAAACATTTCCGATATCGCATT 1675
|||||
Db 26178 AAGTGCTTACTTCTTTTGTGACCTACAAACCAATTTTAAACATTTCCGATATCGCATT 26237
QY 1676 TTTCAACCATGCTCATCAAGACAGTAAGATAAAACATTTTAAACATTTTAAACATTTTAA 1735
26238 TTTCAACCATGCTCATCAAGACAGTAAGATAAAACATTTTAAACATTTTAAACATTTTAA 26297
QY 1736 AACCATCTGCTCGTAGGAATGCTTATTTTCTTCTACGACGAATATATACCCGCGCTTTT 1795
Db 26298 AACCATCTGCTCGTAGGAATGCTTATTTTCTTCTACGACGAATATATACCCGCGCTTTT 26357
QY 1796 CAATTAACACTAACTCCACATATAGTAACCCCTTAATTTTAAATAACCGCAATTTA 1855
26358 CAATTAACACTAACTCCACATATAGTAACCCCTTAATTTTAAATAACCGCAATTTA 26417
Db 1856 TTTGGCGGCAACACAGGATCTCTTTTAAAGTTACTCTCTATTAACATACGTTTCCATCT 1915
26418 TTTGGCGGCAACACAGGATCTCTTTTAAAGTTACTCTCTATTAACATACGTTTCCATCT 26477
QY 1916 AAAAATTAGTAGTATGAACTTAAACGGGGCATCGTATTTAGTTTCCATATTTAGCTTT 1975
Db 26478 AAAAATTAGTAGTATGAACTTAAACGGGGCATCGTATTTAGTTTCCATATTTAGCTTT 26537
QY 1976 CTGCTTCCTTTTGATAACCACTGTTATTCATGTTGCAATGCTGACCTGTTATACCAAC 2035
Db 26538 CTGCTTCCTTTTGATAACCACTGTTATTCATGTTGCAATGCTGACCTGTTATACCAAC 26597
QY 2036 GATATAGTCTATTAATGCAATATAGTATCGCCGAACGATTAAGCTCTTACAGGCTTCTGAA 2095
Db 26598 GATATAGTCTATTAATGCAATATAGTATCGCCGAACGATTAAGCTCTTACAGGCTTCTGAA 26657
QY 2096 GAAAGCTTCAAGTACTAATTAAGCCGATAGTACGACCGAGCTTCTGAGCAATTTTCAT 2155
|||||

Db 26658 GAAGCGTTTCAAGTACTAATAAGCCGATAGATAGCCAGCACTTCGTAGCCATTTCAT 26717

Qy 2156 AAGTGTAACTTCGCGCTCCGCTCATPAACAGACATTCAGTATGCGGAAAGT 2215

Db 26718 AAGTGTAACTTCGCGCTCCGCTCATPAACAGACATTCAGTATGCGGAAAGT 26777

Qy 2216 ATGCATGCTGGGTGTGGGAAGTGTGAAGAAAGAAAGTACGCTGCGTGTGACATC 2275

Db 26778 ATGCATGCTGGGTGTGGGAAGTGTGAAGAAAGAAAGTACGCTGCGTGTGACATC 26837

Qy 2276 ACTGCTATCTTCTTACTGTTATGCAAGTCTGAGTGGGTGGACACA 2322

Db 26838 ACTGCTATCTTCTTACTGTTATGCAAGTCTGAGTGGGTGGACACA 26884

RESULT 9

RSSTR1

LOCUS R.serpentina str1 gene for strictosidine synthase. 2242 bp DNA linear PLN 22-JUN-1992

DEFINITION X62334.1 GI:21126

ACCESSION X62334

VERSION str1 gene; strictosidine synthase.

KEYWORDS Rauvolfia serpentina (serpentwood)

SOURCE Rauvolfia serpentina

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Gentianales; Apocynaceae; Rauvolfioideae; Vinceae; Rauvolfia.

REFERENCE 1 (bases 1 to 2242)

AUTHORS Bracher,D. and Kutchan,T.M.

TITLE Strictosidine synthase from Rauvolfia serpentina: analysis of a gene involved in indole alkaloid biosynthesis

JOURNAL Arch. Biochem. Biophys. 294 (2), 717-723 (1992)

MEDLINE 92231602

PUBMED 1567228

REFERENCE 2 (bases 1 to 2242)

AUTHORS Kutchan,T.M.

TITLE Direct Submission

JOURNAL Submitted (27-NOV-1991) T.M. Kutchan, Lehrstuhl fuer Pharmazeutische Biologie, Universitaet Muenchen, Karlstr 29, D-W-8000 Muenchen 2

FEATURES

source

1. 2242

/organism="Rauvolfia serpentina"

/mol_type="genomic DNA"

/db_xref="taxon:4060"

/cell_line="cell suspension culture"

/clone_lib="lambda charon 4A"

320. 330

434. 440

466. 1581

/gene="str1"

/prim_transcript 466. >1581

/gene="str1"

/evidence=experimental

547. 1581

/gene="str1"

/codon_start=1

/product="strictosidine synthase"

/protein_id="CAA44208.1"

/db_xref="GI:21127"

/db_xref="GOA:P15324"

/db_xref="SWISS-PROT:P15324"

/translation="MAKLSDSQTMALFTVLLFLSSSLASSPILKEILIEAPSYAPN SFTFDSTNKGFTSYVDGRVIKYEGPNSGFVDFAYASPYWNKAFCENSTDAEKPLCG RTYDISYNLQNNQLYIVDCYYHL.SVVGSEGHATOLATSDVGVPEKWL.YAVTDQRTG IVYFTDVSTLYDRGVQOIMDTSKTRLIKYDPSKETLLKELHVPGEAEVSADS SFVLVAEFLSHQIVKVMLEGPCKGTAEVLYKIPNPNIKRNADGHPWSSSELDGNM HGRVDPKGIKPFDEFGNILVILPLPPFAGHHEQIQEHDLGLYIGILFHGSVGLIVYD KKGNSFVSSH"

polyA_signal 1672. 1676

polyA_signal 1691. 1695

misc_signal 1797. 1804

ORIGIN /note="GT box"

Query Match 20.9%; Score 486.2; DB 8; Length 2242;

Best Local Similarity 97.3%; Pred. No. 1.8e-91;

Matches 505; Conservative 0; Mismatches 13; Indels 1; Gaps 1;

Qy 927 ACAACCAATAAATGCTGAATGAATTTCTAAGCGGAGATCGCTAGTATTTAACTATT 986

Db 1703 AAAAAAAAAAAAAAAAAAAGAAATCTAAGCGGAGATCGCTAGTATTTAACTATT 1762

Qy 987 GCTGCAGCATTTCTGAGTCCATATATAAGTATTTGTAACCTTTTGTGTCAGGTTG 1046

Db 1763 GCTGCAGCATTTCTGAGTCCATATATAAGTATTTGTAACCTTTTGTGTCAGGTTG 1822

Qy 1047 TTCTTTAGAGAGTAAAGATCAATGCACTAAACGAACTGAACCAAGCGATCGAAA 1106

Db 1823 TTCTTTAGAGAGTAAAGATCAATGCACTAAACGAACTGAACCAAGCGATCGAAA 1882

Qy 1107 ATATCCCTTTGGGATTTCTGACTCGATAAGTCTATTAATTTTCAGAGAAAAATATTCATT 1166

Db 1883 ATATCCCTTT-GGATTTCTGACTCGATAAGTCTATTAATTTTCAGAGAAAAATATTCATT 1941

Qy 1167 GTTTCTGGGTGGTGGATTGCAACCAATCATTCATTCAAAATTGTTGTTTACACACCC 1226

Db 1942 GTTTCTGGGTGGTGGATTGCAACCAATCATTCATTCAAAATTGTTGTTTACACACCC 2001

Qy 1227 ATTCGCGCCGATAAAGCATGAATGTTGTCGTCGGGCA TAGAATTAACCGTCACCTGAAA 1286

Db 2002 ATTCGCGCCGATAAAGCATGAATGTTGTCGTCGGGCA TAGAATTAACCGTCACCTGAAA 2061

Qy 1287 AGGTATAGTTAAATCACTGAATCCCGGAGCACTTTTCTATTAATGAAGAGTGAATC 1346

Db 2062 AGGTATAGTTAAATCACTGAATCCCGGAGCACTTTTCTATTAATGAAGAGTGAATC 2121

Qy 1347 TGACAATTTCTGGCAACCATTTTAACACACGTCGGAATGTCATGAATTTCTGAAGAGT 1406

Db 2122 TGACAATTTCTGGCAACCATTTTAACACACGTCGGAATGTCATGAATTTCTGAAGAGT 2181

Qy 1407 TACCCCTCTAAGTAATGAGTGTTAAGACGCTTTTCATT 1445

Db 2182 TACCCCTCTAAGTAATGAGTGTTAAGACGCTTTTCATT 2220

RESULT 10

AE006063/c

LOCUS AE006063 10124 bp DNA linear BCT 08-MAR-2001

DEFINITION Pasteurella multocida PM70 section 30 of 204 of the complete genome.

ACCESSION AE006063 AE004439

VERSION AE006063.1 GI:12720510

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 10124)

May,B.J., Zhang,Q., Li,L.L., Paustian,M.L., Whittam,T.S. and Kapur,V.

TITLE Complete genomic sequence of Pasteurella multocida, PM70

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (6), 3460-3465 (2001)

PUBMED 11248100

REFERENCE

AUTHORS

2 (bases 1 to 10124)

Zhang,Q. and Kapur,V.

TITLE Direct Submission

JOURNAL Submitted (24-OCT-2000) Department of Veterinary Pathobiology, University of Minnesota, 1971 Commonwealth Ave., St. Paul, MN 55108, USA

FEATURES

source

1. 10124

/organism="Pasteurella multocida"

/mol_type="genomic DNA"

gene /strain="PM70"
/db_xref="taxon:747"
78. .458
/gene="PM0279"
78. .458
/gene="PM0279"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK02363.1"
/db_xref="GI:12720511"
/translation="MSNPVSWFEIHVENLQRAKTEYQAVFQPLPLTDGDCPSGKMSVC
FESTDYAQYGIIGMLYEDKDFPVVRGNNTLFFRACEDCTVEAERAVKFGKLLQGKTS
IGDEGFYAVIEDSEGNRIGLSMQ"
563. .1729
/gene="succ"
/note="synonym: PM0280"
563. .1729
/gene="succ"
/codon_start=1
/transl_table=11
/product="Succ"
/protein_id="AAK02364.1"
/db_xref="GI:12720512"
/translation="MNLHEYQAKQIFAEYQLPVGKGYACKSADAEADAIKKLINGDVWV
AKQVHAGRGKAGVKLVRENAEVRAFADQMLGORLVTPQTDVNGQPVNPLLYEBS
SIARELYLGAVLDRASQRVFMVSTEGVNIEVAEKPPLHKMAIDPLGMPYQG
RELAFKLGKGDQIKQFAHIFVQAKMFVEKDALLVENPLVVTKEGNLLCLDAKIV
DSNALYROPALKAMQDPSQEDPREALAESHOINVALEGNICMVNAGLAMGTMDII
KLHGQPANFLDVGGATKERVAEAFKIILSDSAVKAVLVNIFGGIVRCDLIAEGIIA
AVNEGVNVPVVRLEGNNAFLGREILANSGVNIIAANTLTDAIQA VKAEGK"
1731. .2600
/gene="succ"
/note="synonym: PM0281"
1731. .2600
/gene="succ"
/codon_start=1
/transl_table=11
/product="Succ"
/protein_id="AAK02365.1"
/db_xref="GI:12720513"
/translation="MSILINKDTKVICQGFSGGTFHSEQALAYGTQLVGGVSPGKG
GTHLGLPVFNIVDAVDATGATASVITYPAPGCKDALLEAIDAGIQLVICTEGIPT
LDMLQVKQRLNQTVRMIGNCPGVITPECKIGIMPGNIHKKGKIGIVRSGLTYE
AVKQTTDEGFGQSTCVGIGDPIPGSNFIDILKLFQDDPQTEAIVMIGEIGSAAEEA
AAFIKDYVTKPVVSYTAGVTAPKGRMGHAGALISGGKGTADKFALEAGVKTIRS
LAEIGALKALK"
2845. .4398
/gene="PM0282"
2845. .4398
/gene="PM0282"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK02366.1"
/db_xref="GI:12720514"
/translation="MRFRLNRDESMPONAQNLVCLLIDNMNDYSYITMFMSFDDSD
GVLHEIGNIKIGFKQVEYETIPTHAEIKNRFDRVFDKLPDNFISLGQDVDFYSRIEFL
SNEVKDII LKNLNDVYVNPDDLDEFONEPVFYSLMRHVSEPTIKEQFRVLEGKSP
TDTFSFIRNRQDSMGKITLFEVKA NSVPSTNIHAII GRNGVGKTTLLNGMISFEIN
KNQYKDGAIFYKKEKIHFFTNPIKISDEYFSRLVAVSFVPDFIPNNDEKNYSYIGL
KKSETQLKEPKKYFEEDFWLSLECRSFSFKRDLWYKALQNLESNNFAEMNLSHLMN
ECFSKDIVLQFINRMSSGHAFLVLTITQLVAKVEKTLVLLDEPESHLPPLISAFIR
TISELJDNRNAIAYATHSFVLQEI PKSCIMKIVRTGIETKAYRLVTETPGNIIGIL
TRVFGLEVEKSGFHLINESVNKGTYESILESINYQLKEGKLLKTLIRLDEKN
QGFFYYDKD"
4358. .5290
/gene="PM0283"
4358. .5290
/gene="PM0283"
/codon_start=1
/transl_table=11
/product="unknown"

gene /protein_id="AAK02367.1"
/db_xref="GI:12720515"
/translation="MKKIKGFIFIMIKILPTDCQNSNIIDSCCNGILRDELTKIIL
VNKGLLRISKYYLKLAKNNCLFKYKXHI CHSDEIIRMGRFSDLLTCTDMKKLYEDY
FVRKRCGSYYDKIINYAKDPLIQCPFCGGIGVPSLDHFLPKVHFAYYSIFPYNLWPI
CKDCNKSCKNEFPYTEINKQILHPYLDDDCFFEEQWLFANYCIENNIASGTNYVYSP
PTNWSEDKKGKVFHEEQFKLDERFSIRAPYLVDLIAQITRAKSKSKLSEADCIIEII
LDPIINSEBYNKNHMKVALYQAVKNKFTHTWNI"
complement (5372. .6016)
/gene="adk"
/note="synonym: PM0284"
complement (5372. .6016)
/gene="adk"
/codon_start=1
/transl_table=11
/product="Adk"
/protein_id="AAK02368.1"
/db_xref="GI:12720516"
/translation="MKIILGAPGAGKGTQAOPI MNKFGIPQISTGDMLRGAIKAGTD
IGKQAKTLMDAQOLVPDDLII SLVKERVAQDCAKGFLLDGFPTIIPQADALKVGIQ
IDYVLEFDPDDEVIVERMSGRVRHQASGRYHVYVNP PKVEGKDDVTGEDLIRADDK
BETVLDRLKVYSTTKPLVDYQAEAKAGNTKXFERLDGTYKVEKVSQELDTIIA"
complement (6165. .7442)
/gene="PM0285"
complement (6165. .7442)
/gene="PM0285"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK02369.1"
/db_xref="GI:12720517"
/translation="METQKPSLLSQIFSRNMLCIFTGFSSGLPLYVLPQWVPAMLSS
SNVDVKTIGFFTLIGFPYTWKFIISPLDRYYBNFLGRRSRMMFMQVALLALLALLG
QYDPVANITYAVIYSTIAVFSAVQDIVIDAYRBEILTDEELGLGNSIHVNAYRISGL
IPGGLSLILADSLPMDVVEIFTALFMI PGLILSLFLSKEPDVPQIDRSLPEYHTPALP
FKEFFSRKGVASALGLIFFI FLKFGDSLATSLOTKFILDMGFSKTDIGYVVKLNSLM
ASII SCMLGGIIMLRIGINRSLMLFGFVQLLTILGFVMMASYGKFNITIGSFEFFVLT
VIIGEYIGVGLGTAAFAVAFMARETNPLYTATQLAIFTSLSAIPSKIFSAFSGVLVADY
GYTFFWICFPFAIPGMLMLFKVAPWMSAEQRA"
complement (7445. .8461)
/gene="galE"
/note="synonym: PM0286"
complement (7445. .8461)
/gene="galE"
/codon_start=1
/transl_table=11
/product="GalE"
/protein_id="AAK02370.1"
/db_xref="GI:12720518"
/translation="MALIVTGAGYIGSHVVELLNANKDQVVLNDLNCNSSPKSLERY
AOITGQVKFYQGDILDTALLQKIFAENQIQSYVIEHAGLKA VESVQKPAEYMNANVT
GSLVLIQEMKAKGAWNPFVSSSATVYGDEPIIPITESCKVGTTNPYGTSKEMVQILL
KDI AKATPEFSITILRYENPVGAHESGLIGEDPNGIPNNLLPYISQVAIGKLPQSVF
GSDYETHDGTGYVDYIHVVDLAIGHLKALDRHESDAGLHIYNLGTGSGYSVLDMYKAF
EKNVDIKIPYKLVDRRPDIATCYSDPSLAKTELMNTAARGLQMMKDTIWHQKNPK
GYRD"
8728. .10041
/gene="PM0287"
8728. .10041
/gene="PM0287"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK02371.1"
/db_xref="GI:12720519"
/translation="MSLEKHPEMTKRGSNVRQEVIALGTTFLIAMVYSVIVENMLGAA
GFPABSVFIATCLVAGLSLLIGLWANVPMALGCAISLTAFATSLVLGQOYSIPVAL
GAVFLMGVFTTISATGIRAMILRNLPSSIAHGAGIGIGFLLLIAANGVGLVYVNSQA
GLPVKMGFTSPVLMSLIGLALIIGLEKMRVGAII MWIIATLVAGLFDPNVKEGG
SIFKMPTEGNSLFLLELDLMGALQPAIIPVFPALVMTAVFDATGTIRAVAGQANLDDK
EGQIINGKALTSDSVSSIFSGIFGTAPAAVYTESAAGTAAGKGTGTTAVVVGVLFL
MLFPQPLAFLVPGYATAPALMYVGLMLSNVSKLDFDDFIGAMSGLVCAVFIVLTANI
VTGIMLGFAALVIGRVVSGEVKKLVNVTGLIAIAIVAFYAFGWA I"

ORIGIN

Query Match		14.2%;	Score 329.8;	DB 1;	Length 10124;
Best Local Similarity		53.7%;	Pred. No. 8.5e-59;		
Matches 857; Conservative		0;	Mismatches 632;	Indels 108;	Gaps 4;
QY	250	TTACCTTCATCTCTGTTCATATCATCGCTTTTAAACGGTTCGACCTTCTAATCCATC	309		
Db	4367	TTTTTTTCATCTCTTAATCTGATAGAGTCTTTAATATTAATTTAACCTTCCTTACCTAAT	4308		
QY	310	TGACCATTAATATTTTGTAGATGGTTTTCATAGAAGCTCTGAATCAACGGAGCTGGCAT	369		
Db	4307	TGTAATTAATACTTCTTAATAAATACTTTCATATGTTCCCTTATTTACGGATTCAATTG	4248		
QY	370	AATAAGTGGTGGTATCCAGAAATTTGTCACTTCAAGTAAGAAACACCTCAGAGTTAAACA	429		
Db	4247	AGTAATTTATGAAAAACCTGATTTTCTACTTCTAATCCAAGACTTCCCTCGTCAAAATA	4188		
QY	430	CCTAAGTTCTCACCCGAATGTCTCAATATCCGACGGATTAATTTATTGCTTCTCTGAC	489		
Db	4187	CCGATATTTTCACCAAAAGTCTCAGTTACAGCCTATATGCTTAGTTTCTATTCCTGTT	4128		
QY	490	CGTAGACTTTCCACATGACAGATTTTGGAACTCTTGACATCTACTGGGAATGAGTT	549		
Db	4127	CGAACATTTTCCAGATACAGATTTTAGGTATTTCCGTGTAACACTACCGGTGATGAGTT	4068		
QY	550	GCAATTAATGCTACACCATTGGCTGCATCGAGTAAGTCGCTTAATGTTGTAAGAAAGCA	609		
Db	4067	GCAACAATTTGCAATAGCATTTGCAATTAATTAATTCAGATTAAGTTGTAATAAGCA	4008		
QY	610	GAGAGCAAAAGGTGATGCAGATGAACCTCTGTTTCATCGAATTAACATAATGACTTTTGG	669		
Db	4007	GATTAATAGCGGAGATGCAAAATGGCTTTCGCTCATCAAGTAAGTAAGCAAGTGTGTTTCT	3948		
QY	670	CGAACGACATCTACTAATCTTGATAGTAATAATAAAACAATGTCATGTCAGAGCTCATTT	729		
Db	3947	TCAACTTTTCGCAACTAATTTGTAATAGTCAATTAAGCAAAAGCATGGCCTGAACCTCATC	3888		
QY	730	CGAAGCAGATATTTCTGGAATTTGTCATAAAACAATTTAGTAATTTATCATCGTCCACT	789		
Db	3887	CGGTTGATAAAGTGCAGTACTAATCTTTACTAAA-----	3853		
QY	790	TGAATCTGTGTTTATTACGTTTAACTCTTCATATTTAGAAATGAGGCTGATGAGTTC	849		
Db	3852	-----ACATTCAATTCATTAATGCGCTTAATTC	3825		
QY	850	ATATTTGAAAAAGTTTTCATCAGTACTAGTTTGTGATAGCTTCAAGCCAGAGTTGCTTT	909		
Db	3824	ATTTCAACAAAATTAATTTGTCAGACTCAAGATTTCTGAATTTGCTTTAACCCCAAGATCTCTT	3765		
QY	910	TTTCTATCTACTCTCATACAAACAATAAATGCTGAATGAATTTCTAAGCGGAGATCGCCT	969		
Db	3764	TTAAAGGAGAAAGATCTACATTTTGAATAAGACAGCCAAAATCTTCTTCAAAATATTTT	3705		
QY	970	AGTATTTTAAACTATTTGCTGGCAGCATTTCTTGAGTCCAAATATAAAGTATTGTACTT	1029		
Db	3704	TTTGGTCTTTTAATTAAGTTTCAAGATTTTAAACCAATATAAGATAATTTT-----	3649		
QY	1030	TTTGCTGGGTCAAGTGTCTTTAGAGAGAGTAAGAGATCAAAATGCACCTAAACGAAACT	1089		
Db	3648	-----TTTCATCGTTAATTCGGAATAAATGATGCAAGACACTAAAGACACG	3603		
QY	1090	GAAACAAGCGATCGAAAATATCCCTTTGGGATCTT-----GACTCG	1131		
Db	3602	GCAACTAAACGGCTAAAAATATTCGTCATAATCTTATTTGAGTATTAGTAAAAAAATGT	3543		
QY	1132	ATAAGTCTATTAATTTTCAGAGAAAAATATTCAATGTTTCTGGGTTGGTGTGATGCACCA	1191		
Db	3542	ATTTTTCCTTCTTATAAAATGCTCCATCTTTATATCATTTTATTTATAAATGACTTA	3483		
QY	1192	ATCATTCGATTCAAAATTTGTTGTTTACACACCCGATTCGCGCGATATAAAGCAGTAATG	1251		
Db	3482	ATCATTCGTTAAGTAGTGTGTTTTCACAAACACCATTTCTTCTATAATTTGCATGATTA	3423		

QY	1252	TTGCTGCTGGCATAGATTAACCGTCACTCAAGAGTATAGTTAAATCAGTGAATCCG	1311		
Db	3422	TTAGTACTAGGAACCTGAATTTGCTTACTTCAATTTCTAGAGTTATTTTCCCCACTACTG	3363		
QY	1312	GGAGCACTTTTCTATTAATGAAGAGTGGAAATCTGCAATTTCTGGCAAAACCATTTAAC	1371		
Db	3362	TCTTGCTATTACGAATTAAGAAAAAGTAAATATCAGTTAATGAGATTTGCCCTTCTTAA	3303		
QY	1372	ACAGTGGCACTGTCCATGAATTTCTGAAAGAGTTACCCCTTAAGTAATGAGGTGTTA	1431		
Db	3302	ACTCTTCTAACTGCTTTAATCGTCGGCTCAGTCAATGCTCATTAACGAGGTATTA	3243		
QY	1432	AGGACGCTTTCATTTCAATGTCGGCTAATCGATTGGCCATACTAATAATCCTGAATA	1491		
Db	3242	AAAAACCGGTTTCAATTTGAATTCATCTAAAGATCAGGGTTAATAACAACATCATTTCAA	3183		
QY	1492	GCTTTAAGAGGTATGTTTAAACCATCGCTTAATTTGCTGAGATTAACATAGTAGTCA	1551		
Db	3182	TTTTTTAAGATTATATCTTTTACTTCATTTAGAAAGTTCAAAAATTTCTGAAATAGAAATCA	3123		
QY	1552	ATGCTTTCACCTAAGAAAAAACATTTCAAGGAGTT-----GACTGAAT	1596		
Db	3122	ACATCTTGACCAAGAAATAAAATTAATCTGTAATTTATCAAAACACTCTCCCATCAAAAT	3063		
QY	1597	TTTTATCTATTATGAATAAGTCTTACTTCTTTTGAACCTACAAACCAATTTTA	1656		
Db	3062	CTGTTTTTAATTTCTGCATGAGTTGGAATAGTTTCAACTGTCTTTAAACCAATCTTA	3003		
QY	1657	ACATTTCCGATATCGCATTTTTCACCATGCTCATCAAGACAGTAAGATAAACAATTTGA	1716		
Db	3002	ATATTCCTATTTTCATGCAAGACTCCATCAGTATCAAGAAACATCATAAACATAGTA	2943		
QY	1717	ACAAAGAAATAGTCATTTCCAACCATCTGCTCGTAGGAATGCTTATTTTCTACTGCA	1776		
Db	2942	ATATTAAGAATATCATTTCCAAATTAATCAATTAATAAACAATAAGATTTGGCATTTCTGG	2883		
QY	1777	GGAATATACCGCGCTCTTTCATAATAACATAACTCCA	1813		
Db	2882	GGCATGCTTTCATCTCTAATTAATAATCTAATTCGCA	2846		

RESULT 11	
AE016869	
LOCUS	310967 bp DNA linear BCT 09-OCT-2003
DEFINITION	Pseudomonas syringae pv. tomato str. DC3000 section 14 of 21 of the complete genome.
ACCESSION	AE016869 AE016853
VERSION	AE016869.1 GI:28854015
KEYWORDS	
SOURCE	
ORGANISM	Pseudomonas syringae pv. tomato str. DC3000
REFERENCE	Pseudomonas syringae pv. tomato str. DC3000
AUTHORS	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas.
1 (bases 1 to 310967)	
Buell, C.R., Joardar, V., Lindeberg, M., Selengut, J., Paulsen, I.T., Gwin, M.L., Dodson, R.J., Deboy, R.T., Durkin, A.S., Kolonay, J.F., Madupu, R., Daugherty, S., Brinkac, L., Beanan, M.J., Haft, D.H., Nelson, W.C., Davidse, T., Zafar, N., Zhou, L., Liu, J., Yuan, Q., Khouri, H., Fedorova, N., Tran, B., Russell, D., Berry, K., Uterback, T., Van Aken, S.E., Feldblyum, T.V., D'Ascenzo, M., Delaney, T.P., Lazarowitz, S.G., Martin, G.B., Schneider, D.J., Tang, X., Bender, C.L., White, O., Fraser, C.M. and Collier, A.	
The complete genome sequence of the Arabidopsis and tomato pathogen Pseudomonas syringae pv. tomato DC3000	
Proc. Natl. Acad. Sci. U.S.A. 100 (18), 10181-10186 (2003)	
12928499	
JOURNAL PUBMED	
REFERENCE	
AUTHORS	
2 (bases 1 to 310967)	
Buell, R., Joardar, V., Khouri, H., Fedorova, N., Tran, B., Russell, D., Berry, K., Uterback, T., Van Aken, S., Feldblyum, T., Gwin, M., Dodson, R., Deboy, R., Durkin, A., Kolonay, J., Madupu, R., Daugherty, S., Brinkac, L., Beanan, M., Haft, D., Selengut, J.,	

TITLE Nelson,W., Davidsen,T., White,O., Fraser,C. and Colmer,A.
JOURNAL Direct Submission
Submitted (03-MAR-2003) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
FEATURES
source
Location/Qualifiers
1. 310967
/organism="Pseudomonas syringae pv. tomato str. DC3000"
/mol_type="genomic DNA"
/strain="DC3000"
/db_xref="taxon:223283"
/note="pathovar: tomato"
/locus_tag="PSPTO3607"
complement(283. 1092)
/locus_tag="PSPTO3607"
/locus_tag="PSPTO3607"
/note="The annotation of this multicopy IS21 family
element is based in part on an HMM match, the ISfinder
database, and on positional information.; similar to
GP:6009454, and GP:17865161; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="ISPsy4, transposition helper protein"
/protein_id="AAO57081.1"
/db_xref="GI:28854016"
/translation="MNLQHARLTCLKGLERGVDPWPHLAQQAASGEDSFADLEK
LLAETDARSESRQALLKTAALPAVKTLQYDFAFATGVPRALQELALSFVGRAE
NIVFLGPSGVGKSHLATALAYRAVMAGIKTRFVTAADIMQLTAHROERLKEYFSRV
VMAFGLLVIDEIGYLPFGRDEANLFNNVAKRYEQSLILTSNLPFTWAGTFADDQT
LTAAMLDRLHHAHIVQMTGESYRLKDKRKAGTKSSRAEPARKYPEGGQN"
complement(1089. 2111)
/locus_tag="PSPTO3608"
complement(1089. 2111)
/locus_tag="PSPTO3608"
/note="The annotation of this multicopy IS21 family
element is based in part on an HMM match, the ISfinder
database, and on positional information.; similar to
GP:14594862, and GP:17865160; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="ISPsy4, transposase"
/protein_id="AAO57082.1"
/db_xref="GI:28854017"
/translation="MLTQEQSVEIKVLARQGHGIFARELGISRNTVRKYLKARS
PSDKVRPARPCKIDPFKDYLERIEARPHWIPATVLLREITALGYSGVSLKAYIR
PFKRKAEEPVRFETLPGQIQLVDETTIRGRQPKAFVATLGFSAFVRSEREDS
EAWLTGLREAFAYFGVPEQALFDNAGTITTERDAFEGQHWRPRLALADEFGFIP
KVCRPYRAOTKGKVERNGYLKGSFITPLATLKSAGLTLDVTANAHIGWMDVAH
QRHGTGTGVPAVRLAQEQVLLPLPTQSLRPQAGLRIGRVLPYESTLQHPLSVEQ
LLEVRA"
2191. 3213
/locus_tag="PSPTO3609"
2191. 3213
/locus_tag="PSPTO3609"
/note="Warning: Expression or function of this gene might
be disrupted. The first 2 residues of the coding sequence
are contributed by the end-sequence of an insertion
sequence, ISPsy4.; identified by match to PFAM protein
family HMM PF00589"
/codon_start=1
/transl_table=11
/product="site-specific recombinase, phage integrase
family"
/protein_id="AAO57083.1"
/db_xref="GI:28854018"
/translation="MTLGEIGSRRSWESTGRALYDFESFLQASHLHWKDVGEEKD
LVTGYRDYCRDTCSLATSTIRQVRVYCEFYDFSLKNKTEKLPYNEVSREYKSRSEN
SSRHPIASKTVLVKDVLVYLTKSLPFLRMSEVKELTHAQNPHRMIRFALHTGMR
QDEIASFTAYVFNPEKGNKTSRNIQIRLDRPDGTGMRTKRSKMRDIFISRRFMLHL
AYLTKMRGERSTLANSTAKELFLNHRGEPYADFGKSCRTIRNIGKSVSTHMLR
HTYATQTLTSLQKNSEIEPLVFLQRLGHSSIQTTMVYLLHVLNALADEAVLAYDELA
NLSEVA"

gene
CDS
3210. .4889
/locus_tag="PSPTO3610"
3210. .4889
/locus_tag="PSPTO3610"
/note="identified by match to PFAM protein family HMM
PF02899"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAO57084.1"
/db_xref="GI:28854019"
/translation="MTGKKTFITDLSIPAVNHNSTENQTLKLTDRFPVTRVDI
GKNATRMWIDFERWYRADIDEIAYACORQIERFVARQDRDIEPSTVASICGGLRHF
LDYVMLRAKAYERELRLADVDRSLIDFLGYLNDLAISRQSRNFTAKSILMALCA
RGLAYVDSGDHATFPKNPEPNTPKTDGSPSPKROQEVAAALKLTAPITWDDVIT
TSSLSYALTTLAHTGRNTTPLLMDRNCIRAHPKDNLTVLWKRGHSTYSKSTLR
SPPQDRALRESIPYVRTNIEQLIRRVHLTGLSDAPEDIANRAMIYRGQASARGK
VVELNDAMLKFALQKLVKEYDLKDTDGKPLQNVSRLRKTFANRIYELGNTQITAI
ALGNTPAVTEQDYLRPTTESKRNWKFGEFLVRELMTKTIGATFHTPVGKCRDDEG
QFAPKSSGGVCFEFLNCVRCKHYAVTADDLTKLFSFYIRIYSESSIGKIRWEKESH
IPRLIDNYIIIEGLRRKIFKPEVEIARTRANTPHFWTCALNSLEAFL"
5227. 5260
/locus_tag="PSPTO3611"
/note="This region contains an authentic frame shift and
is not the result of a sequencing artifact; This gene is
disrupted by an IS element.; IS52, transposase,
interruption-N"
5343. 5702
/locus_tag="PSPTO3612"
5343. 5702
/locus_tag="PSPTO3612"
/codon_start=1
/transl_table=11
/product="ISPsy5, Orf1"
/protein_id="AAO57085.1"
/db_xref="GI:28854020"
/translation="MMRPDAKVEKVLYPEKPYDERKSIDGLAALVELDIKAVEDPVL
FVFLNKSRRNRVKILTYWERNGFCLWLKRLSEBRPKISPDVDEAIVLVQELNWLIDGF
DLMRNRPHKILTPREVA"
5720. 7273
/locus_tag="PSPTO3613"
5720. 7273
/locus_tag="PSPTO3613"
/note="similar to GP:8037785; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="ISPsy5, transposase"
/protein_id="AAO57086.1"
/db_xref="GI:28854021"
/translation="MRPLPDNLPPDDPVLTKOMLIESYEARSREQEVKDAYATHIVDLK
EQIKLRLDRIFGRKSQVTEPNTPOLALFNEPSEPMPLVGDADAEVVAAPARRGKX
PLSADLPRIEVIHELBEHELTCACGCRKHVSEETSEQDIVPMQIRVIXKIRKYGC
RGCEAPVTAGKPAQILIEKSMASPSVLAMLLTTRYVDGLPLHREFTVLSRHGIELPRQ
TLARWIIQCGEHLQPLNLIMRDRLLSPVTHODETRVQVLKEPDRDPTSQSMWVQAS
GPPDHKVLFDYTSRRAQDVPLRLLESYRGYVTTDDYAGYNALALQPGVERLACMAHA
RKFEVAQKVQPKRTGRADIALTMINKLYGERELKDGSDQRFTRGQEKSLPILAQ
LKSWLDXTQSHVTPQSVLGKAVGYLANNWSRLKRYEAGFLPIDNNPAERAIKPEVIG
RKAWLFSDTPEKGATASAIQIYSLVETAKVNGEPTWLRHVLERLPLAOSVADYEALLP
WNCSPEMPR"
7321. 8271
/locus_tag="PSPTO3614"
/note="This region contains an authentic frame shift and
is not the result of a sequencing artifact; This gene is
disrupted by an IS element.; IS52, transposase,
interruption-C; similar to GP:8885851, and GP:1196750;
identified by sequence similarity; putative"
8452. 10053
/locus_tag="PSPTO3615"
8452. 10053
/locus_tag="PSPTO3615"
/note="identified by match to PFAM protein family HMM
PF00589"

gene
CDS

gene

gene

CDS

gene

CDS

gene

gene

CDS


```
/codon_start=1
/transl_table=11
/product="site-specific recombinase, phage integrase
family"
/protein_id="AA057087.1"
/db_xref="GI:28854022"
/translation="MPGQKPSASTVRNLFENFLPFVHYLKSIIQITSMASVTLACTS
VIQTCYSIIQINRYKGPLSASSLTRFAAVEAIYELSQYTADPMRHPMNGAASI
ARASSALANSTSGKPIIPDDVFCALFSAAHQNI TKGSWLDLRDELDSVALSLQHV
FRMTQSAONRLITERGWIIGLVAENLALDLRTSCYIVLASVSGCRNHELNLKTSR
HSTEGDDGTVYHMLRLETOKTSGIHDMVPELAIITHELLEWAVPYQRIEAINL
RTIQNPADPLTAEMLRHDSLFATMSDNIRVIRLTLGSTMENLKRFSRAGLSWDL
ATHQFRKFPANYAAHSFEDLRYLROHFTHWSMNMTLGYAMDQDMGQHFIDLDFEIE
SELGDIKDTVSQMLESDLSISGGYKSLGRWKDPNNLTLFKDHSMWKAIABSTAIR
SNGHAWCTADDDRCVGNSTERTRCGGCTSSVGAQHASVSSAADQLIDLQYDDIGE
GGRRVLRDLERCKTVLLELYIQE"
complement(11292..12278)
/locus_tag="PSP103616"
complement(11292..12278)
/locus_tag="PSP103616"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA057088.1"
```

Query Match 8.7%; Score 201.4; DB 1; Length 310967;
Best Local Similarity 47.7%; Pred. No. 4.5e-32;
Matches 730; Conservative 0; Mismatches 751; Indels 48; Gaps 3;

```
QY 262 CTGTTCAATATCATCGCTTTTAAACGGTTGCACTTCTAATCTCATGACCATTAATA 321
Db 92237 CTGCTGGCGACCATCTTTCGCAAGATTGCTGTGCTCAACCCGAGCTGCTCTGATAG 92296
QY 322 TTTTGTAGATGTTTCATAGAAGAACTCTGAATCAACGAGCTGCATATAAGTGTGG 381
Db 92297 TCGGCTAAATACTTTCAATGTCCCTCCAGATTGACAGCTCTTGTACACACTCGTGG 92356
QY 382 TATCCAGAATTTGTCACTTCAAGTAAACACCTCAGAGTTAAACACCTAAGTTCTCA 441
Db 92357 TAAACCGATTATATGACTTCCAGCCAAATATTTCCGGGTAAGAACAACCCGATTTCTCA 92416
QY 442 CCGAATGCTCATATCCGACGGATAATATTATTGCTTCTTGAACCGTAGACTTTC 501
Db 92417 CCGAAGCTTTCCCTCTCAGGTCGATCAAGCTCTGCTTCAAGCTCGCAGTCCGGTGAGCTTC 92476
QY 502 CACATGCAGGATTTTGAACCTCTTGCAAGTACTACTGGGGAATGAGTTGCAATTATTGCT 561
Db 92477 CACACACAAGAACCCCGCGCTCTTGAACAACACGCGGAGTGAAGTTGCGATAATTGCA 92536
QY 562 ACACCAATGCGTCATCGAAGTGAAGTCTTAATGTTCTGTAAGAAAAGCAGAGAGCAAGGT 621
Db 92537 ACACCGTTCCGCTCTGTAGCAGGTCAAGAAATAGCTCGGTTGAACGCGACAGTAAAGGG 92596
QY 622 GGATGCAGATGAACCTCTGTTTCATCGAATAAACTAATGACTTTTCGCCAAGCACTCT 681
Db 92597 GGGTGAATGACTTTCAAGCTCATCCATTAGCACAAGCGTTTTCCTCGACGGCTCT 92656
QY 682 ACTAATCTTGTAGTAATAAACAATTCATGCTCCAGAGCTCATTCGAGACAGATAT 741
Db 92657 ACAAGCTTTGTCAATGTCAACAATACAATGAGTGCCTGAACTCATCTGCTAATAATA 92716
QY 742 TTCTGATATTTGCTATAAACAAATTTAGTAATTTATCATGCTCCACTGAATCTGTGGT 801
Db 92717 CGATGCGCTTAAACAATAGCTCGTCCGTGGGAAGCTCAAT-----TAA 92761
QY 802 TCATTACGCTTAACCTTTCATATTTAGAAATGAGGCTGATGAGTTCCATATTTGAAAAG 861
Db 92762 TGAAGCAGGTCATGTCTTCGAAATGCTGTGATTCGAGACGCTTGATCGACGCAAC 92821
QY 862 TTTTCATCACTACTAGTTTTTTGATAGCTTCAAGCCAGAGTGTCTTTTCTATCTACT 921
Db 92822 CATCTTTGCCGCTTGCTTCCGACTAAAGCAAGACTCCAGACTTGTATATAAGTCTTCC 92881
```

```
QY 922 CTCATACAACCAATTAATGCTGAATGAATTCTTAAGCGAGATCGCTAGTATTTTAA 981
Db 92882 GTCAATGCCCTCAGTCTGTTGGGGGGCGCTTACCAGGGCCTTCCCGGGGATTGCAATG 92941
QY 982 CTATTGCTGGCAGCATTTCTGAGTCCAATATAAAGTAATTGTACCTTTTGCTGGTCA 1041
Db 92942 -----CCACGTAATAATAGGACACACCTTTACTTCTATCG 92977
QY 1042 GGTGTTCTTTAGAGAGTAATAAGATCAATATGCACTTAACGAACTGAAACAGCCGAT 1101
Db 92978 GGTGATCGGCAGGAACAAGATGATCAACCGCTTAATAAGACACAGAAACACACTG 93037
QY 1102 CGAAATATCCCTTTGGGATTTCTGACTCGATAAGTCTATTATTTTCAGAGAAAAATAT 1161
Db 93038 CTAAATTAATCCGTCGGCATGGGTTCACTGCCAGCGAATCATAGATCAAAAAATTTCCG 93097
QY 1162 TCATTGTTTCTGGTGGT-----GATTGCAACCATTCATTCCAAATTTGTT 1212
Db 93098 CAATGGGCTCTGCCCGTGTATCGCTGAGAATGTTTCAGCCATGTCATTATATTTGTT 93157
QY 1213 GTTTTACCACACCCATTCGCCCGCATTAAGCATGAATGTCTGCTGGCATAGAAATTA 1272
Db 93158 GTCTTGCTACACCGTTCGCCCGCATGAGTACATGAACATTCGTCGAGGCAAGAAATGG 93217
QY 1273 ACCGTCACTCAAAAGGTATAGTTAATCACTGAATCCGGAGACACTTTTCTATTAAT 1332
Db 93218 GGGCTTACTTTAAACGACACGTCACGCCCGCGCTGCTCATCCCATTTGTCTGTGTA 93277
QY 1333 GAAAAGTGAATCTGCAATTTCTGGCAAAACCATTTTAACAACGTCGCAACTGTCATGA 1392
Db 93278 CTAAATGAAAACCTGTGACTCTGCTTCCCGGCCAGACACGCTTGAACGTGTACCG 93337
QY 1393 ATTTCTGAAAAGATTACCCCTTAAGTAATGAGGTGTTAAGGACGCTTCAATTTCAATG 1452
Db 93338 ATGACAGACACGCTGACATTTCCGATCAGCAGCTCGCAACAACCTGCTCATTTCTTAACG 93397
QY 1453 TCGGCTAATGATTTGGCCATACTAATAATCCTGAATAGCTTTAAGAAAGTTAAGTTTA 1512
Db 93398 TCCTCAATCGCAGGGGTGGCGATTAACATCTTTGAGCGCTTCAAGACAGCAACGCT 93457
QY 1513 AAACCATCGCTTAATTTGCTGAGATTAACATAGTAGTCAATGCTTTCACCTAAGAAAA 1572
Db 93458 GCGTCGGCCGACAAACCAACCTGAAGTTTCTATTAATAATCGACATCCTGACCAAGCAA 93517
QY 1573 AACATTCAGGAGGTGACTGAATTTTATCTATTAATGAATAAGTGTACTTCTTCT 1632
Db 93518 AAAAAATCATCACTCAGGCTTAATTTTGTGTAATACTTCTCGGACGTCATCCGTGC 93577
QY 1633 TTTGACCTACAAAACCAATTTTAACATTTCCGATATCGCATTTTTCACCATGCTCATCA 1692
Db 93578 TGCTGTCCTTATATCAATCTTGTGTGTGTCGCAAGCTCAAGTTTTTTGAAGCTCGGACCG 93637
QY 1693 AAGACAGTAAGATAAACAATTGTAAACAAGGAATGATTCACCAACATCTGCTGTAAG 1752
Db 93638 AAGTCATGACTGTGAACCTGTTTGAATGAGTAGTGTCCACGCACTTCCAAACCAAC 93697
QY 1753 AATGCTTATTTTCTTCTACTGCAAGAAAT 1781
Db 93698 AGCAAGTATCTGTTAAGTCAGACGGTAT 93726
```

RESULT 12
ML0672112/c 150050 bp DNA linear BCT 02-SEP-2002
LOCUS ML0672112
DEFINITION Mesorhizobium loti R7A symbiosis island; segment 1/4.
ACCESSION AL672112 AL672111
VERSION AL672112.1 GI:20803827
KEYWORDS
SOURCE Mesorhizobium loti
ORGANISM Mesorhizobium loti
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Phyllobacteriaceae; Mesorhizobium.

REFERENCE 1 (bases 1 to 150050)
AUTHORS Sullivan,J.T., Trzebiatowski,J.R., Cruickshank,R.W., Gouzy,J., Brown,S.D., Elliot,R.M., Fleetwood,D.J., McCallum,N.G., Rossbach,U., Stuart,G.S., Weaver,J.E., Webby,R.J., de Bruijn,F.J. and Ronson,C.W.
TITLE Comparative sequence analysis of the symbiosis island of Mesorhizobium loti strain R7A
JOURNAL J. Bacteriol. 184 (11), 3086-3095 (2002)
MEDLINE 2199272
PUBMED 12003951
REFERENCE 2 (bases 1 to 150050)
AUTHORS Ronson,C.W.
TITLE Direct Submission
JOURNAL Submitted (07-FEB-2002)
COMMENT Department of Microbiology, University of Otago, P.O. Box 56, Dunedin, New Zealand; MSU-DOE Plant Research Laboratory, Michigan State University, East Lansing, Michigan 48824 USA; Laboratoire de Biologie Moleculaire des Relations Plantes-Microorganismes, Unite Mixte de Recherche (UMR) 215 Centre National de la Recherche Scientifique (CNRS), Institut National de la Recherche Agronomique, Chemin de Borde Rouge, BP27, F-31326 Castanet Tolosan Cedex, France. clive.ronson@stonebow.otago.ac.nz
http://sequence.toulouse.inra.fr/msi.
location/Qualifiers
source 1.150050
/organism="Mesorhizobium loti"
/mol_type="genomic DNA"
/strain="R7A"
/db_xref="taxon:381"
198..1457
/gene="ints"
198..1457
/gene="ints"
/function="elements of external origin; phage-related functions and prophages"
/note="Gene name confidence : probable"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="P4-TYPE SYMBIOSIS ISLAND INTEGRASE PROTEIN"
/protein_id="CAD31406.1"
/db_xref="GI:20803828"
/db_xref="GOA:O68526"
/db_xref="SPTREMBL:O68526"
/translation="MEVPKMAISDVKCRNARPISKI VKLSDGGGLQLWVQPTGSRIMR LAYRFGKQKLALGSYPLISLAERARDAKRLLRIGIDPAQERKSQKASAEATER SIAEDYVDKLKKEGRADRTITKVKLDFAYPTFGKSVREIDPATILVALRSVEVRG RYESARLRSTIGSVFRYAIAATADVPTIALRGALVGPTVPRAVTDPRKLGLL RAINAFDGPPTTRALKLMALFPRPGLRAEWDEFDESASVMSIPEGRMKRRPHR VPLSRQAVSVLTSLEISDGALVFPVSRSRSPISDNTINALRRMGYKKEETAHG FRATASTLINECGKWHPDATERQLAHIESNDVRAYARAHEWERIKMMQMWADYLDE LESKDAGVISIGHGKSTVHTPSRDLR"
complement(1845..3002)
/gene="msi002"
complement(1845..3002)
/gene="msi002"
/function="miscellaneous; unknown"
/note="Product confidence : hypothetical
Gene name confidence : hypothetical"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="HYPOTHETICAL PROTEIN"
/protein_id="CAD31407.1"
/db_xref="GI:20803829"
/db_xref="SPTREMBL:O68527"
/translation="MRWDPSLTLDSDFTMFWRRRISAKRKVLVIGRGFVLTLET ARRLHALDADISVQLVFDNGLGDSPKRSELTATNHEGLLKLVPAEKTIVAPIEIGP VGAAVTPRNTKASLDRAQDCRDFDDVIDVISAPRWALTAVSVLLFKLDKLAENGPSVNLVHTSVESVADLGAAGSLRDSVTFVLGFGSHLEQTQNMPRVWFPVLGEGQR DRLHLIDHVNPDICPVLPEPTRGRGDEVGHRDTLFDDEKVEPANILASEYN PFEAYROIPTAMERYKALKPMGGCKLFVSPSSKLLSIGVLLACYDHKFGKSSDRF DVGIPVETAVYGDPEQNAETADELHLSLWIRGEWEPIPATASAAVTAAAT"

gene
complement(3016..4908)
/gene="msi003"
complement(3016..4908)
/gene="msi003"
/function="miscellaneous; unknown"
/note="Product confidence : hypothetical
Gene name confidence : hypothetical"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="HYPOTHETICAL PROTEIN"
/protein_id="CAD31408.1"
/db_xref="GI:20803830"
/db_xref="SPTREMBL:Q8KH03"
/translation="MRASQRSTSDDEFVRLFGVGALEVIKSFDDPWNGLVPLRSAPGG KTTFLRLTPRPLKLTNLDADQHQVKATRDALREVGAIGRDGPKLGMVYFTSEY KELSADFNGNLRRELLNSRIVATLRSVLERAERQEPDDLDOYSFEWEPESDATIPG RATGKELFTWASRIERDFYDRMDLGS PASIKG HARLDGLKWFARHVRISDAQGMDV KRVLMDDELQTLAPPOQASLIEFITNAREQGWIAERLEALT HKDLISEGALRKRDY ENVQLEERWSGTRTKSYTKFVETIANLRAKADGFEGRLEPLI GETDIPSL ETAL TGRTAIEIAEVYKMGAVGSPRYSNWTIANARGLSGNLLERAFKWLKILVTRDTRTQT SEFDSLTLTDEFARSGSGTDRAAEHFLRTDVGAPIYFGRSLSAVSSSNVQYLEVA GALFEEISAKIRFHRDQVPPLTAARQDALIRGVAKDRWDGLPRLPRGVEARLLLEAI GEYCRQOTFRSSAPYAPGVGTGIAITMDRKLIDSEDAEISHFIRLRDVMISLVSHNL LMPRLDHRNNGREYVVFYLNRLLCVHFGPLPGYGGWRHQTLSLHMQEAAGAKAIAVA EEA TLV"
complement(4927..6045)
/gene="msi004"
complement(4927..6045)
/gene="msi004"
/function="miscellaneous; unknown"
/note="Product confidence : hypothetical
Gene name confidence : hypothetical"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="HYPOTHETICAL PROTEIN"
/protein_id="CAD31409.1"
/db_xref="GI:20803831"
/db_xref="SPTREMBL:Q8KH02"
/translation="MRKDLAESLAKIMEMSDBEKAERAYLIESFASYKDYEQFSP GKRFIESLALWLGQFTGGEEREAAYNFRKRLIFISTDEMNHLVELFTPIIRPILIA DTATLEGVDPHKXKALVDTEVEYRTRLRQTLVGLSDGARLDMFRANPQEMSNQVEH AYDVSDPKSEGWODLRKHLTKIIGREPHEARFRIYVLDDFTGSGTSFIREDCQG GWTGKIATIVSGLMKAGNLGDALIANSGVKIILILYVAGDQAI EHI ESR LPKLTFKGT VEEFEVHKLG ETTPLDDASDSPILGLVADDRYFDDADDEHSKVGTSKRYGFANCR L PVLAHNTPNNISYILMAEDDQSVRGLFPVRSRHKLQ"
complement(6149..6520)
/gene="msi005"
complement(6149..6520)
/gene="msi005"
/function="miscellaneous; unknown"
/note="Product confidence : hypothetical
Gene name confidence : hypothetical"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="HYPOTHETICAL PROTEIN"
/protein_id="CAD31410.1"
/db_xref="GI:20803832"
/db_xref="SPTREMBL:Q8KH01"
/translation="MQIYMKPKQIERAELVRHINERKLEAGLSYAEIADIADVASQV SRICRGQFITFGASVRICTTGLQNTQGBGTTWRSRHAADPNWAKLERSIKRAWDN TPGAERLAKVIAAVGEITRK"
6544..7041
/gene="msi006"
6544..7041
/gene="msi006"
/function="miscellaneous; hypothetical/global homology"
/note="Product confidence : hypothetical
Gene name confidence : hypothetical"
/codon_start=1
/evidence=not experimental

```

/transl_table=11
/product="HYPOTHEITICAL TRANSMEMBRANE PROTEIN"
/protein_id="CAD31411.1"
/db_xref="GI:20803833"
/db_xref="SPTREMBL:Q8KH00"
/translation="MEFFVTDVGQEIFRYRASEKPYGAFSNLFRREIERGERFATSE
HAYOAGKARKPEVKAMLMMAASPALLAMAHGLYYWDVSPGMSRTKFDNRKCVLRKAF
TQHEDLRELLVSTGSARLIETATVDNEVNLWGVNGQGRNMLGELLMEIRAKLQND
VSKAA"
gene
complement(7080..7988)
/gene="msi007"
complement(7080..7988)
CDS
/gene="msi007"
/function="miscellaneous; hypothetical/partial homology"
/note="Product confidence : hypothetical
Gene name confidence : hypothetical
predicted by Framed"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="HYPOTHEITICAL PROTEIN"
/protein_id="CAD31412.1"
/db_xref="GI:20803834"
/db_xref="SPTREMBL:Q8KGZ9"
/translation="MSAVDVVGTFVLDRIYADSLAVSSEDLGSGCNVLISLMLQ
RNVAPVLSLGEDEVGIRLVAFRHAGAEIRYIRQTEHIRSPIAQRHGGVHTFEFR
CPETSIDLPRPEPIDMEDVSRALPAIKNCNRYTDLRSLGAIVDAMEADAGSTVYFE
PSALGDKELFRRAVAAASVLKFSARLATALAGVAPHDALFLIVTHGEAGLEVRGAE
KAWCKSYPAPELRDTCGSGDMVSVGVLDLISRRDRATRKISIDEFVVGIEAGQRLAA
ENCAYVVGARGIFRERGAIFYARSVLRG"
complement(8041..9972)
gene
complement(8041..9972)
CDS
/gene="msi008"
complement(8041..9972)
/gene="msi008"
/function="miscellaneous; unknown"
/note="Product confidence : hypothetical
Gene name confidence : hypothetical"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="HYPOTHEITICAL TRANSMEMBRANE PROTEIN"
/protein_id="CAD31413.1"
/db_xref="GI:20803835"
/db_xref="SPTREMBL:Q8KGZ8"
/translation="WTVFSEKLRRLPETIELASAHGADIAPFALSQGRRLAVAIASG

Query Match      6.1%; Score 142.2; DB 1; Length 150050;
Best Local Similarity 57.9%; Pred. No. 1.2e-19;
Matches 252; Conservative 0; Mismatches 183; Indels 0; Gaps 0;

QY      293 GACCTTCTAATCCTATCTGACCATTAATTTTGTAGATGGTTTCATAAGAAAGCTCTG 352
Db      101499 GGCCTTCCAAGCCAACCCGACCAAGCAATGCTCGAGATGCTCTTCAAGTCTCGACCT 101440
QY      353 AATCAACGGACTGCGATAATAAGTGTGTATCCAGAAATTGTCACTTCAAGTAAAAACA 412
Db      101439 CCTCCACCTTGTCCCGCAACATCTTGTGGAAGCCCGAATCGGTGACCTCCAGCCGAGA 101380
QY      413 CCTCAGAGTTAAAAACACCTAAGTTCTCACCGGAATGTCATATATCCGACGGAATATAT 472
Db      101379 TTTCCTGTGTCAGTATTCGACGTTTCTCCAAAGTTTCAATCTCGGCTCTCCGCGA 101320
QY      473 TTATGTCTTCTTGACCCGTAAGACTTTCACACATGCAGAGATTTTGAACCTCTTGACAGTA 532
Db      101319 TCACAGTTCGTCATTTGCTGAACTTTCAGACGCACTCATTTGGCACTCTCTGAGAA 101260
QY      533 CTACTGGGAATGAGTTGCAATTATTTGCTACACCATTCGCTGCATCGAGTAAGTCGCTTA 592
Db      101259 CGACAGGCGAGTGTGTTGCGATGATCGACGCCGCTTGCGGTTGATCAGAGAGTCGAAA 101200
QY      593 ATGTTCTGTAATAAAGCAGAGCAAAAGGTGATGCAGATGAACCTCTGTTTCATGCAATA 652
Db      101199 GGGCTCGACAAATGCGACATATAGCGGTGATGAGATGCGCTTCAGGTTCTGTCAGAA 101140
```

```

QY      653 AAATAATGACTTTTCGCCAAGCACATCTACTAATCTTGTGATAGTAATAACAATTG 712
Db      101139 GAACGAGTGTCTCTCTCTCAACGGTTTCCACCAGACGGGTAATGTCAGTAGAACGATTT 101080
QY      713 CATGTCAGAGCTCA 727
Db      101079 TGTGACCCGAGCTGA 101065

RESULT 13
BX572601/c
LOCUS
DEFINITION
Rhodopsudomonas palustris CGA009 complete genome; segment 9/16.
ACCESSION
BX572601 BX571963
VERSION
BX572601.1 GI:39649375
KEYWORDS
complete genomes.
SOURCE
Rhodopsudomonas palustris CGA009
ORGANISM
Rhodopsudomonas palustris CGA009
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Bradyrhizobiaceae; Rhodopsudomonas.
REFERENCE
1 (bases 1 to 348580)
Larimer, F.W., Chain, P., Hauser, L., Lamerdin, J., Malfatti, S., Do, L.,
Land, M.L., Pelletier, D.A., Beatty, T.J., Lang, A.S., Tabita, F.R.,
Gibson, J.L., Hanson, T.E., Torres y Torres, J., Peres, C.,
Harrison, F.H., Gibson, J., and Harwood, C.S.
Complete genome sequence of the metabolically versatile
photosynthetic bacterium Rhodopsudomonas palustris
Unpublished
2 (bases 1 to 348580)
Larimer, F.W. and Harwood, C.S.
Rhodopsudomonas genome consortium
Submitted (24-JUL-2003) Submitted on behalf of the Rhodopsudomonas
genome consortium, the DOE Joint Genome Institute, Production
Genomics Facility, 2800 Mitchell Drive, Walnut Creek, CA 94598,
USA, and the Genome Analysis Group, Oak Ridge National Laboratory,
1060 Commerce Park Drive, Oak Ridge, TN 37831, USA;
larimerf@ornl.gov

FEATURES
source
1. 348580
/organism="Rhodopsudomonas palustris CGA009"
/mol_type="genomic DNA"
/strain="CGA009"
/db_xref="taxon:258594"
194..3916
/locus_tag="RPA2457"
194..3916
/locus_tag="RPA2457"
/locus_tag="RPA2457"
/function="InterPro IPR002048"
/note="observed by proteomics
citation: Proteomics from Verberkmoes et al. (2003)
unpublished"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="CAE27898.1"
/db_xref="GI:39649376"
/translation="MPQERILRDGAHPSREEPAPDARVQREAMAKPTPDQSDLAAG
HDCNDPHMHEGHMDHEHEEDARHARKLLKRNGLGHRIDDWPSRRTPGAPWLKRTLL
TIAVGAVLIAGFAGLMLRLGVGINVDWVTPWLADAIEENIGQNTVEIGTQIERA
GRIVALRLRDIIVVRDRDHVVASAPKAEVRLSGGALLTGKLRAESLRVLGAEALARI
DPDGQTVVASGSTDRPLATGIGAPSGRPAERKLPQATPDAGAVAPATPSGSPTPP
PPATDTSDTMKGVLAGLEWLDLSLSGLDGNLINEIGLKNGLVVDQQRNRTFEN
ISLSLRPSGTGVALSVGEEGGLKAWSLRYQVPAHDGVRSVELHANQVPTRTLALL
RLKNLTYSADPELTDLKGVEGRDGVPTYFRGKLTAGAGTVIDSDTPDYEMAIQIET
SFEMDANRRVLIAPFKVLAGPNRVTLASLEPPNGSVNDWRLGLSGGTIVLPGAQNES
PLIFNRIAVVRPDSANRRVLITQADVSNGEIGVAGSGSIDYSGEPRLTGLAGTPMS
VSALKRMNPALIVPELRWVNDRLAGSIQTIDIAINSPMKNLSRGRPIPDGGLLVN
ILNGGAIRPVDGMPVVKDADMRVRVTRTATVAINQANADTDGGRVALSDILFEVP
DLAPKPARIKFKLEGPVPAVAEVLASGRLSDAALPIDPNSSKENVSAOVALGMPI
QRELTKQDTSYSITADLNGFSADKLANMQKLEANTLKVIANNQGYQYKGVKINGQAA
```



```

misc_feature
/locus_tag="RPA2457"
/note="1 probable transmembrane helix predicted
byTMHMM2.0"
complement(3886..3907)
/note="TransTerm prediction, 98%"
complement(4065..4625)
/locus_tag="RPA2458"
complement(4065..4625)
/locus_tag="RPA2458"
/function="InterPro IPR000516 COGS COG3038"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="putative cytochrome b561"
/protein_id="CAE27899.1"
/db_xref="GI:39649377"
/translation="MTTTLDTADKGAQYALAVLIHWLIFRAVIALFASMNVAHGL
EKTDLRRTLYDWHKAVGTAVGLAVLRILWIKIHGAPDLVSPITEVIARISHGL
YLLLAVPITGLGMTLFGSGRDLLGIPPLAKNDIAGALHAHEHLFLTAIVLIIH
VIGALWQYFRHDATLGRMPWLRGK"
order(complement(4500..4565), complement(4398..4454),
complement(4269..4334), complement(4116..4181))
/locus_tag="RPA2458"
/note="4 probable transmembrane helices predicted
byTMHMM2.0"
complement(4846..6063)
/locus_tag="RPA2459"
complement(4846..6063)
/locus_tag="RPA2459"
/function="InterPro IPR007114 COGS COG0477"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="putative permease"
/protein_id="CAE27900.1"
/db_xref="GI:39649378"
/translation="MDHRASRTGFIILGLCFVLSLLGRGLGESFTVFLPISGSFGWD
RAEVSYISLTALCTGLASPLVGLLFDPSGPRVYTLGMTLLGALLIAVAERLWL
QLTVGLCVGLGIAFTGNVNSILILGRMFGARLPTAMAVVYSAIGAVLMLPIAQLLI
DRYGMREAYLILIGAMLILVPLSIMPWRFAAGADAAORPAODEDGTILGSAMR
HHAFWALFATFFFTALIGMYSIAQVAYLVDAFTPLQATAMGFSGVLLVGMGLVS
WLDGVTGRPSILFSYASTIGVILMLLOWYPNLLLGFTVCFGSMIGSRGLITA
TAMSIFRGRRVGTIYGTISISGSLGSAFGSWGGLLHDLTQSYNPVLGFALVSVLGM
IPFLVPALRERO"
complement(5980..6063)
/locus_tag="RPA2459"
/note="Signal predicted by SignalP 2.0 HMM (signal
peptideprobability 0.601) with cleavage site probability
0.299 atresidue 28"
order(complement(5980..6036), complement(5854..5919),
complement(5782..5832), complement(5686..5751),
complement(5584..5649), complement(5488..5553),
complement(5397), complement(5221..5286),
complement(5149..5199), complement(5068..5133),
complement(4972..5028), complement(4861..4926))
/locus_tag="RPA2459"
/note="12 probable transmembrane helices predicted
byTMHMM2.0"
complement(6148..7404)
/locus_tag="RPA2460"
complement(6148..7404)
/locus_tag="RPA2460"
/locus_tag="RPA2460"

```

Query Match	6.1%; Score 142.2; DB 1; Length 348580;
Best local Similarity	57.0%; Pred. No. 1.1e-19;
Matches 280; Conservative	0; Mismatches 208; Indels 3; Gaps 1;

244 TGTACTTTACCTTCATCTCTGTTCAATATCATCGCTTTTAAACGGTTTCGACCTTCTAAT 303

|||||

Db 122447 TATAGTTTTTCGCGTCGCGATCAGCAATTAGCGCCCGAAGATGCGCGCTTCCGCA 122388

QY 304 CCTATCTGACCATTTATATTTTGAAGTGTTCATAGAAAGCTC--TGAATCAACG 360

Db 122387 CCGAGTTGGCCGCCAAAATGTTCCACACACTGCGTCGATTCAGACCGGAGTAGTACG 122328

QY 361 GACTGGCATATATAGTGTGTGATCCAGAAATTTGTCACTTCAAGTAAACACACCTCAGCA 420

Db 122327 GCGTCATTCAACAAGCTATGGAACCAAGACTTTGTGACCTCAAAAGCCGAACTTCCCGG 122268

QY 421 GTTAAACACCTAAGTTCTCACCAGATGTCTCAATATCCGACGGATAATATTATTGCT 480

Db 122267 GTCAAGATGCCGACATTTCTACCAAGGTCTCGACGGTTGGCGCTCGACACGGCGACC 122208

QY 481 TCTCTGACCGTAGGACTTTCACATGACAGATTGGAACCTCTTGACGACTACTGCGG 540

Db 122207 GCGCGCGACGCGGAGTTTCCAGCACAGAGCTTAGGCACTTCTGAAAGCACTACCGGT 122148

QY 541 GAATGAGTTGCAATTATGCTACACCATTTGCGTGCATCGAGTAAAGTCCCTAATGTTGCT 600

Db 122147 GAATGCGTCGCAACGATCGCCACACCGTTGCGTTGACGAGAAGATCGACAGCGCCGA 122088

QY 601 AAAAAAGCAGAGAGCAAGGTGATGACAGATGAACCTCTGTTTCATCGAATAAACTAAT 660

Db 122087 ATGAAAGCCGAGAGACGCGGGGATGAAGATGCCCTCGGGCTCTCCAGTAGGACCAAT 122028

QY 661 GACTTTTCGCCAAGCAGACATCTACTAATCTTGTGATAGTAAATAAACAAATTCATGTCGA 720

Db 122027 GTGCGTTCTGTCAGCACCTTCAAGAGGCGGCTCATCGTTAGAAAGAACAAATGCGTCC 121968

QY 721 GAGCTCATTCG 731

Db 121967 GAACTCAATCG 121957

RESULT 14

SLHAUR/c

LOCUS SLHAUR 2365 bp DNA linear BCT 09-JAN-1995

DEFINITION S. lividans haur gene.

ACCESSION X77574

VERSION X77574.1 GI:455520

KEYWORDS ORF; resistance protein.

SOURCE Streptomyces lividans

ORGANISM Streptomyces lividans

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycetaceae; Streptomyces.

REFERENCE 1

AUTHORS Zhou, X., Deng, Z., Hopwood, D.A. and Kieser, T.

TITLE Streptomyces lividans 66 contains a gene for phage resistance which is similar to the phage lambda eas9 endonuclease gene

JOURNAL Mol. Microbiol. 12 (5), 789-797 (1994)

MEDLINE 94328930

PUBMED 8052130

REFERENCE 2 (bases 1 to 2365)

AUTHORS Kieser, T.

TITLE Direct Submission

JOURNAL Submitted (07-FEB-1994) T. Kieser, John Innes Institute, Colney Lane, Norwich NR4 7UH, UK

FEATURES

source

1..2365

/organism="Streptomyces lividans"

/mol_type="genomic DNA"

/strain="66"

/db_xref="taxon:1916"

1035..2358

/gene="haur"

1039..2358

/gene="haur"

/codon_start=1

/transl_table=11

/product="phage Haur3 resistance protein"

/protein_id="CAA54681.1"

/db_xref="GI:455521"

/db_xref="GOA:Q54346"

/db_xref="SPRMBL:Q54346"

/translation="MRCHNGSRSVKVAVIGIFEGMEGRTGLPERFEALPPDFSLG

ADEDYVATLRDEDEETRLTILSLGRDVAYYPLEFQALSEPVRDSLRTDTVR

TOYRRIVAGGPILSRFHIRYQEAQSGSAPPLTLELKVDPKPNPSNIHAYISNGVG

KTVLHKIAHTLNSRNRRSSSWLEDRDLHPFTNVVVSFSAFDSHTPHQVNRAY

NRVGDVDYQYVGLKTDDEEVKNYALGAEFARCOVERCEDHPAKARWGVYLAKE

ETDPLFHDLEIRELQAODRDPDRVAFAGSSGKIVLTLARLVQHTTERTLVIDE

PEALHHPPLISTFVRTLSELRLDRNGLAVIATHSPVLQETPREAVWALRRAGDDLRY

DHPEIETFGENVITREIFGLEVRRTGFNRLIQLLADEGYVFRNP"

ORIGIN

Query Match 4.5%; Score 104.8; DB 1; Length 2365;

Best Local Similarity 53.7%; Pred. No. 1.2e-11;

Matches 217; Conservative 0; Mismatches 187; Indels 0; Gaps 0;

QY 329 GAATGTTTCATAAGAAAGCTCTGAATCAACGCACTGCGATTAATAAGTGTGATCCAG 388

Db 2353 GATTTCTCGAAAGACATACCTTCGTCGCGGAGCACTGGATCAACCGGTTAAACCTG 2294

QY 389 AATTGTCATTCAGTAATAAACACCTCAGAGTTAAACACCTAAGTCTCACCGAATG 448

Db 2293 TGCGCGCACTCCCAATCCAAAGATCTCCCGGTGATCACTCCGACGTTCTCGCCGAATG 2234

QY 449 TCTCAATATCCGACCGAATAATATTATTGCTTCTCTTGAACCGTAGACTTTCACATGC 508

Db 2233 TCTGATCTCGGGGTGATCCACACGAGTCACTCGCGCTCTCGGAGGGCCAGACGG 2174

QY 509 AGGATTTGGAACCTCTTGCACTACTGCGGAATGAGTTCGAATTATTGCTACACCAT 568

Db 2173 CTTACAGAGCGCTCTCTGGAAGAACGACTGGGAGTGGATGAGTAACCGCAAGCGGT 2114

QY 569 TCGGTGCATCGAGTAAGTCCTTAATGTTGTAATAAAGCAGAGCAAGAGGTGATGCA 628

Db 2113 TTCGGTCTCTCAGCAGCTCAGACAACGTCGACGACGATCGACAGCAACGCGGGTGA 2054

QY 629 GATGAACCTCTGTTCAATGAATAAACTAATGACTTTCCGCCAAGACATCTACTAATC 688

Db 2053 GATGGGCTCCGGTTCATGATCAGAGCAGAGGCTCGGTGTTGTTGATCAGGCC 1994

QY 689 TTGTGATAGTAATAAACAAATTCATGTCAGAGCTCAATTGCA 732

Db 1993 GGGCCAGAGTTAGCAGACGATCTTGTGTCCGGAATCAGCCCA 1950

RESULT 15

AR204117 10523 bp DNA linear PAT 20-JUN-2002

LOCUS AR204117

DEFINITION Sequence 13 from patent US 6365723.

ACCESSION AR204117

VERSION AR204117.1 GI:21500681

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 10523)

AUTHORS Blattner, F.R., Burland, V., Perna, N.T., Plunkett, G. and Welch, R.

TITLE Sequences of E. coli O157

JOURNAL Patent: US 6365723-A 13 02-APR-2002;

FEATURES

source

1..10523

/organism="unknown"

/mol_type="unassigned DNA"

ORIGIN

Query Match 3.2%; Score 75; DB 6; Length 10523;

Best Local Similarity 80.4%; Pred. No. 1.8e-05;

Matches 115; Conservative 0; Mismatches 20; Indels 8; Gaps 2;

QY 2168 CCGCTCCTCGCTCATACAGACATTCACACTACAGTTATGCGGAAAGGTATGCATGCTGG 2227

Db 10258 CCAATCCTTGCTCATACAGGCAATTCACCTTCAGTCATGGCAAGAAAGTATGCATGCTGGA 10317

QY	2228	TTGGGGGAGTCGTGAAGAAGAGAGTCAGCTGCCGTTTGACATCACTGCTATCTTC	2287
Db	10318	-----GAACTCGTGAAGAAGAGAGAGTCAGCTGCCGTTTGCTCAGCTTATCTTC	10371
QY	2288	TTACTGGTTATGCAGGTCGTAGT	2310
Db	10372	--ATTGGCTATGCAAGTCGTAAAT	10392

Search completed: October 6, 2004, 07:57:31
Job time : 9055 secs